DOCTORAL THESIS

Doctorate with International Award

Psychological factors implicated in decision making about health and health outcomes

Doctoral Candidate:

Dafina Petrova

Thesis Supervisors:

Dr. Rocío Garcia-Retamero

Dr. Andrés Catena

Dr. Edward Cokely

Department of Experimental Psychology

Mind, Brain, and Behavior Research Center (CIMCYC)

University of Granada

April 2016

Editor: Universidad de Granada. Tesis Doctorales
Autora: Dafina Georgieva Petrova
ISBN: 978-84-9125-917-6
URI: http://hdl.handle.net/10481/43833

Cover designed by Dobromir Strezov

The research included in this thesis has been sponsored by the scholarship program for research personnel training of the Spanish Ministry of Economy and Competitiveness (BES-2012-051811) as part of the projects PSI2011-22954 and PSI2014-51842-R.

El presente trabajo de investigación ha sido financiado a través del programa de becas para la formación de personal investigador del Ministerio de Economía y Competitividad (BES-2012-051811) asociado a los proyectos (PSI2011-22954 PSI2014-51842-R).

Health is one of those things we truly value only once we lose them. I dedicate this work to everyone who has come to know the true value of health.

I also dedicate this work to my parents who paved my way to the greatest privileges – knowledge and education.

CONTENTS

Contents

EXTENDED SUMMARY
IN SPANISH

FACTORES PSICOLÓGICOS IMPLICADOS EN LA SALUD Y EN LA TOMA DE DECISIONES SOBRE LA SALUD

Resumen en español

Hoy en día muchos pacientes y médicos han de tomar decisiones basadas en información numérica sobre los posibles beneficios y daños de los tratamientos sanitarios (p. ej., la probabilidad de sufrir un infarto tomando estatinas). Sin embargo, tanto los pacientes como sus médicos presentan dificultades a la hora de comprender esta información. Asimismo, muchas decisiones sobre la salud se ven influidas por factores más allá de la evidencia numérica, como las emociones (p. ej., miedo al cáncer) y el conocimiento previo (p. ej., las expectativas). Para promover la salud y la toma de decisiones informada necesitamos identificar aquellos factores que facilitan o impiden la realización de decisiones adecuadas.

En la presente tesis doctoral se propone un modelo teórico general sobre cómo diferentes factores psicológicos influyen en la salud y la toma de decisiones. Los objetivos principales que se han perseguido han sido (1) identificar grupos vulnerables con necesidades de intervención y (2) ayudar a diseñar intervenciones efectivas que proporcionen información adecuada para facilitar la toma de decisiones informada y mejorar la salud de estos grupos. Entre los factores que hemos investigado se incluyen: el formato o modo en el que se proporciona la información, las habilidades cognitivas, las emociones, las creencias y expectativas, y el nivel de apoyo social recibido.

En este trabajo, se ha utilizado un enfoque interdisciplinar. En concreto, se han empleado teorías y métodos propios de diversas áreas, entre las que se incluye la comunicación de riesgos, los juicios y la toma de decisiones, la psicología de la salud, la epidemiología, y la medicina. Asimismo, se han empleado dos metodologías científicas principalmente, los estudios

experimentales y la investigación de campo. En particular, se han llevado a cabo y se han analizado seis experimentos con pacientes y médicos, dos estudios transversales con pacientes hospitalizados, una encuesta nacional representativa, un meta-análisis, y un estudio cualitativo. La muestra de participantes incluye pacientes y médicos procedentes de diferentes países (Estados Unidos, Reino Unido, Países Bajos, y España). Además, se han seleccionado diversos problemas de salud prioritarios como el cáncer, el virus del Ebola, las enfermedades coronarias y las enfermedades de transmisión sexual, y se ha registrado un amplio abanico de variables dependientes (p. ej., el nivel de comprensión y las estimaciones sobre los riesgos y los beneficios de varios tratamientos, intenciones conductuales, decisiones, conducta, y gravedad de la enfermedad en pacientes).

Los resultados de este conjunto de estudios coinciden en indicar que proporcionar información relevante, transparente, y fiable puede mejorar la salud influyendo en la toma de decisiones informada. Los pacientes que comprenden la información numérica sobre los beneficios y los riesgos toman decisiones mejores y desean participar más activamente en la toma de decisiones. Asimismo, los médicos que comprenden mejor la información numérica proporcionan a sus pacientes una información más completa y adecuada sobre los posibles beneficios y riesgos de las opciones disponibles.

Los resultados de esta tesis tienen implicaciones para la promoción de la salud y la toma de decisiones informada. En concreto, plantean sugerencias que pueden mejorar el diseño de las intervenciones y los materiales empleados para comunicar riesgos médicos, especialmente aquellos destinados a los grupos desfavorecidos identificados en este trabajo (p. ej., pacientes con bajas habilidades numéricas y sin apoyo social). Entre las implicaciones teóricas de esta tesis podemos destacar la influencia a nivel teórico de la comprensión de la información numérica sobre la calidad de las decisiones de alto riesgo. Para poder aplicar los modelos tradicionales de comportamiento en salud a situaciones donde se recomienda la toma de decisiones informada, estos modelos deberían incorporar la comprensión de información numérica sobre beneficios y riesgos.

En resumen, las bajas habilidades numéricas, las emociones, y la falta de apoyo social pueden impedir la toma de decisiones adecuada y pueden tener efectos negativos importantes sobre la salud. Sin embargo, hemos estudiado procedimientos novedosos (p. ej., el uso de apoyos visuales), que pueden facilitar la comprensión y la comunicación. Estos procedimientos pueden ser efectivos incluso cuando las decisiones tienen connotaciones emocionales fuertes, o cuando la información que se proporciona es muy compleja o contraintuitiva. La comprensión de riesgos y beneficios a su vez no sólo mejora las decisiones, sino también ayuda a las personas a entender que sus preferencias y valores son de gran importancia en algunas decisiones. De esa manera, la comprensión facilita la toma de decisiones informada y compartida entre pacientes y médicos, aumentando la satisfacción de los pacientes, reduciendo las ineficiencias en el sistema de salud y mejorando la salud pública.

THESIS PUBLICATIONS

Publications included in the thesis

Petrova, D., & Garcia-Retamero, R. (2015). Effective evidence-based programs for preventing sexually-transmitted diseases: A meta-analysis. *Current HIV Research*, *13*(5), 432-438.

Petrova, D., Garcia-Retamero, R., & Catena, A. (2015). Lonely hearts don't get checked: On the role of social support in screening for cardiovascular risk. *Preventive Medicine*, *81*, 202-208.

Petrova, D., Garcia-Retamero, R., Catena, A., & van der Pligt, J. (in press). To screen or not to screen: What factors influence complex screening decisions? *Journal of Experimental Psychology: Applied.*

Petrova, D., Garcia-Retamero, R., & Cokely, E. (2015). Understanding harms and benefits of cancer screening: a model of factors that shape informed decision making. *Medical Decision Making*, *35*(7), 847-858.

Petrova, D., Garcia-Retamero, R., & van der Pligt, J. (in press). What factors moderate self-other discrepancies in decisions? Results from a vaccination scenario. *The Spanish Journal of Psychology.*

Petrova, D., Gray Brunton, C., Jaeger, M., Lenneis, A., Muñoz, R., Garcia-Retamero, R., Todorova, I. (2015). The views of young women on HPV vaccine communication in four European countries. *Current HIV Research*, *13*(5), 347-358.

Garcia-Retamero, R., Petrova, D., Arrebola Moreno, A., Catena, A., & Ramírez Hernández, J. A. (accepted). Type D personality is related to severity of acute coronary syndrome. *British Journal of Health Psychology.*

Cokely, E. T., Petrova, D., Feltz, A., Ripberger, J. T., Gigerenzer, G., & Garcia-Retamero, R. (under review). Informed decision making during a public health crisis.

Petrova, D., Garcia-Retamero, R., Delaney, B., Cokely, E., & Kostopoulou, O. (under review). Strengths and gaps in physicians' risk communication about screening: the role of numeracy.

Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E., Heredia Carrasco, A., Arrebola Moreno, A., & Ramírez Hernández, J. A. (under review). Numeracy predicts risk of pre-hospital decision delay: A retrospective study of acute coronary syndrome survival.

Related publications and sources

Petrova, D., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making*, *27*(3), 191-199. (Master Thesis)

Gray Brunton, C., Farver, I., Jaeger, M., Lenneis, A., Parve, K., Patarcic, D., Petrova, D., Hogg, R., Kennedy, C., Garcia-Retamero, R., & Todorova, I. (2014). Young women's constructions of the HPV vaccine: a cross-cultural, qualitative study in Scotland, Spain, Serbia and Bulgaria. *International Journal of Behavioral Medicine, 21*(1), 11-19.

Science for Informed Decision Making at **www.riskliteracy.org**: A nonprofit university-based project designed to help increase awareness about risk literacy (i.e., the ability to understand, evaluate, and make good decisions about risk).

Note: Each chapter of the thesis consists of one article that either has been published or is in the process of being published in a scientific journal. The content presented here may not fully coincide with the final published articles.

PREFACE

In May 2013 Angelina Jolie, a popular actress, director, and humanitarian, published an open letter in the New York Times called "My Medical Choice" (Jolie, A., 2013). Parts of the letter stated the following:

> "…the truth is I carry a "faulty" gene, BRCA1, which sharply increases my risk of developing breast cancer and ovarian cancer. My doctors estimated that I had an 87 percent risk of breast cancer and a 50 percent risk of ovarian cancer, although the risk is different in the case of each woman.
>
> … Once I knew that this was my reality, I decided to be proactive and to minimize the risk as much I could. I made a decision to have a preventive double mastectomy…
>
> Cancer is still a word that strikes fear into people's hearts, producing a deep sense of powerlessness… I wanted to write this to tell other women that the decision to have a mastectomy was not easy. But it is one I am very happy that I made. My chances of developing breast cancer have dropped from 87 percent to under 5 percent. I can tell my children that they don't need to fear they will lose me to breast cancer… I am fortunate to have a partner, Brad Pitt, who is so loving and supportive. So to anyone who has a wife or girlfriend going through this, know that you are a very important part of the transition…
>
> … Life comes with many challenges. The ones that should not scare us are the ones we can take on and take control of."
>
> Angelina Jolie, May 14, 2013, in The New York Times.

While intuition might say that stories about celebrities belong in lifestyle magazines, there is a very good reason why this story is part of a doctoral thesis. Ms. Jolie's letter is an excellent demonstration of the type of difficult health decisions that many people need to face, and the decisions this thesis has aimed to inform. These are decisions about health that are often based on numerical information about risk, and that can often have serious, life-changing consequences. It is also an excellent demonstration of some psychological factors that can influence decisions about health in profound ways, as will be shown in this thesis. One of these factors is numeracy – the cognitive ability to evaluate complex probabilistic information about health, helping people more genuinely understand what it means to have an 87% risk of breast cancer. Another factor is emotion and strong affective responses that are inevitable when the stakes are high, like the fear and powerlessness that cancer often instills. Yet another factor is the social support from others, like that of a loving and supportive husband that may in ways known and unknown to us keep disease at a safe distance.

Medical choices like the one described by Angelina Jolie have and will become more and more common (Coulter & Collins, 2011; National Institute for Clinical Excellence, 2012). While in the not so distant past health and illness were thought to be in the hands of the gods, today we recognize that we can take some control of our health. Taking control means making decisions about health: what to eat, whether to get vaccinated or screened, when to go to the emergency room, etc. While we may make some of these decisions "in the blind", we often have some information about the potential consequences of others. This information can range from simple instructions about how to protect oneself from disease (e.g., how and when to use a condom) to more complex numerical risk and benefit information expressed in probabilistic and technical terms (e.g., side effects from drugs and risks and benefits of cancer screening). One may expect that the latter only concerns medical professionals, whose job is to practice evidence-based medicine. However, the recently introduced practice of shared decision making has shifted some of this responsibility to patients (Salzburg Global Seminar, 2011). Hence, patients often have to make tough decisions informed by complex numerical information. This requires them to be *risk literate,* that is, able to understand, evaluate, and make good decisions about risks and benefits (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Cokely &

Kelley, 2009; Garcia-Retamero & Cokely, 2013; Ghazal, Cokely, & Garcia-Retamero, 2014; Gigerenzer, 2012)

Involving the patient in decision making is especially recommended for "complex" decisions like Angelina Jolie's, where there are several options (e.g., preventive mastectomy vs. drug therapy), each option has important benefits and risks, and the values and opinion of the patient are essential to the decision (Barry & Edgman-Levitan, 2012). That is, *informed decision making* is recommended. This means that the patient understands the relevant benefits, risks, and limitations, considers her preferences, makes a decision consistent with these preferences, and participates in decision making to the extent that she wants (Fowler, Levin, & Sepucha, 2011; Rimer, Briss, Zeller, Chan, & Woolf, 2004). In contrast to decisions like Angelina Jolie's, there are also "simpler" decisions about health. These are decisions, for which there is well-justified expert agreement and evidence-based recommendations regarding which is the desired option. In such cases, besides providing information, it may sometimes be ethically justified to use strategies that encourage people to make a decision or adopt a behavior that is considered to be mostly or exclusively beneficial (e.g., using condoms, quitting smoking, see Bandura, 2004; Betsch et al., 2015; DiClemente, Crosby, & Kegler, 2009). This thesis aimed to be informative for both "complex" and "simple" decisions about health.

An essential part of health promotion and informed decision making is communicating to patients representative and transparent information that facilitates the decision making process (Fowler et al., 2011; Salzburg Global Seminar, 2011). Here, Ms. Jolie's letter is again a striking example of how influential such risk communication can be and how important it is that it is well-intended, well-designed, and thorough. After the release of Ms. Jolie´s story, a significant increase of referrals and preventive bilateral risk-reducing mastectomies was observed in the United Kingdom (Evans et al., 2014). This increase was not only observed among carriers of the "faulty" BRCA1 gene; actually, it was even greater among those without genetic mutations (Evans et al., 2014; 2015). While these results show that the story certainly has increased awareness, this increased awareness has not translated into improved understanding of the risk of breast cancer with and without the mutation (Borzekowski, Guan, Smith, Erby, & Roter, 2013). A representative survey of the United States population showed that exposure to the story appeared to confuse instead of clarify the understanding of the relationship between positive family history and increased cancer risk, and fewer than 10% of respondents could accurately interpret Ms Jolie's risk of developing cancer relative to a woman unaffected by the genetic mutation (Borzekowski et al., 2013). An analysis of the media coverage of the story in the United States, United Kingdom, and Canadian media indicated that such misperceptions may have been fueled by the media (Kamenova, Reshef, & Caulfield, 2013). For example, while the media depicted Ms. Jolie´s decision as courageous and used emotive language, articles frequently left out important medical information like the rarity of her conditions or the importance of evidence-based assessment of preventive options to reduce breast cancer risk (Kamenova et al., 2013).

The aftermath of Ms. Jolie's story illustrates some of the challenges of educating the public about the benefits and risks associated with different medical treatments and healthy behaviors. In this thesis we have aimed to address several such challenges contributing to fundamental theory in the science for informed decision making. For instance, when numbers are complicated and emotions run high, the ideal of informed decision making may be difficult to achieve. Recent research in decision making about health has shown that many individuals, even those with a college education, have difficulties grasping a host of numerical concepts necessary to make informed decision about health (i.e., they have low numeracy) (Anderson &

Schulkin, 2014; Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Garcia-Retamero & Galesic, 2013; Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008). Sometimes the evidence may be so complicated that even physicians themselves may have trouble understanding it and conveying it to their patients (Garcia-Retamero, Wicki, Cokely, & Hanson, 2014; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz & Woloshin, 2007; Wegwarth, Schwartz, Woloshin, Gaissmaier, & Gigerenzer, 2012; Garcia-Retamero et al., 2014). In addition, strong emotions are rarely absent when life and well-being may be at stake. Such emotions can influence people's decisions in profound ways (Ferrer, Klein, Lerner, Reyna & Keltner, 2014, Finucane, Alkahami, Slovic, & Johnson, 2000; Loewenstein & Lerner, 2003; Loewenstein, Weber, Hsee, & Welch, 2001; Slovic, Finucane, Peters, & MacGregor, 2004; Slovic & Peters, 2006). On one hand, emotions can help us understand what is important (Peters, Lipkus, & Diefenback, 2006), but they can also cloud our judgment and bias us in our decisions (Slovic et al., 2004; Petrova, van der Pligt, & Garcia-Retamero, 2014), or even paralyze us when we need to take action (McKinley, Moser, & Dracup, 2000).

Finally, even when no complex information or strong emotion is involved, and the good decision is rather straightforward (e.g., put on a condom, reduce cholesterol) individuals may still fail to follow recommendations. This may be because of beliefs or attitudes that are influencing their decisions (Fishbein, 1979; Rosenstock, 1974; van Steenkiste et al., 2004). This could be also because they lack the necessary knowledge and skills to complete the behavior (Bandura, 2004; Fisher & Fisher, 2009) or they lack social support (Holt-Lunstad, Smith, & Layton, 2010) all of which can be crucial for health decisions and outcomes.

Predicting how this multitude of factors influence decision making can help us improve the public's health. This could mean successfully designing and tailoring risk communication in ways that maximize understanding and informed choice (Salzburg Global Seminar, 2011; Trevena et al., 2013). Decision aids in the form of pamphlets or videos that describe the available options and help patients consider their importance have been found to improve patients' understanding, reduce difficulty, and increase participation in the decision making process (O'Conner et al., 2009). For example, properly designed and tailored simple visual aids can dramatically increase comprehension, and thus indirectly improve health (Garcia-Retamero & Cokely, 2011, 2013, 2014; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012; 2015). Accurately predicting how the variety of relevant cognitive, emotional, and social factors influence decision making about health could also bring us closer to identifying vulnerable groups (e.g., patients who have difficulty understanding important information or following health recommendations). We can then design cost-effective interventions that provide the correct skills and support to make an informed decision or stick to a recommended health behavior (e.g., Nieuwlaat et al., 2014; Johnson, Carey, Marsh, Levin, & Scott-Sheldon, 2003). Building a knowledge base for such interventions has been an overarching goal of this work.

Goal and Scope of the Thesis

The primary goal of the thesis was to investigate how several psychological factors influence decision making about health and health outcomes, in order to make recommendations about how to promote health and facilitate informed decision making among diverse and vulnerable people. To fulfill this goal we have adopted an interdisciplinary approach using theory and knowledge from the areas of risk communication, judgment and decision making, social psychology, cognitive science, health psychology, human factors engeneering, epidemiology, and medicine. To examine the influence of the selected factors, we have conducted both experimental and field research. Chapters 1 to 5 present a total of six studies, in which numerical information about health outcomes was communicated to participants (patients and physicians). In these studies we have used experimental methodologies and investigated how participants used this numerical information, what factors influenced their comprehension, and how the various psychological factors, alone or in combination, influenced their decisions. Chapters 6 to 10 present results of field research. In these chapters we have aimed to validate and extend some of the results obtained in the experiments in naturalistic settings. In addition, we have investigated the influence of psychological factors on health behavior and outcomes directly and in an ecological setting. To estimate generalizability, a diverse set of health contexts was used. The work also strategically investigated health decisions and outcomes that are among the major priorities of the World Health Organization (see www.who.int), including cancer (Chapters 1 to 3), the Ebola virus (Chapter 4), sexually transmitted diseases (Chapters 5, 9, 10), and cardiovascular disease (Chapters 6 to 8).

We investigated the influence of the following factors: information about health, often of numeric probabilistic nature (Chapters 1 to 5, 9, 10), the cognitive skills and representations required to understand such information (Chapters 2 to 6), the format in which this information is presented (Chapters 1 and 4), the emotional responses that may have implications for health and health behavior (Chapters 1, 2, 4, 5, 7), people’s previous beliefs or attitudes towards health behaviors (Chapters 2 and 5), and social support (Chapter 8).

Overview of the Thesis Chapters

Section I: Experimental Research

Chapter 1 includes a model of informed decision making derived from a decision context where complex and often counter-intuitive information needs to be communicated. In this chapter, we present the results of two ecological risk communication experiments with diverse United States residents on the topic of cancer screening. We examine the influence of visual aids, risk comprehension, and emotions on decisions about cancer screenings with benefits and harms. Finally, we identify factors that can improve high-stakes decisions, as well as factors that can interfere with informed decision making.

Chapter 2 replicates and expands the findings presented in Chapter 1 to offer a more complete picture of factors that can affect informed decision making. We report the results of a laboratory experiment in the Netherlands simulating a patient decision about cancer screening. The experiement examines the influence of numeracy, science literacy, emotions, and previous beliefs on comprehension and decisions.

Chapter 3 extends the previous findings to another relevant population, medical professionals. We present the results of a risk communication experiment with medical professionals in the United Kingdom on the topic of cancer screening. The experiment shows how physicians' numeracy influences their risk understanding. It also examines how physicians' numeracy, patients' numeracy, and the presence of an official guideline about screening influences the quality of risk communication offered by physicians.

Chapter 4 extends the previous findings to a different context – that of emotionally charged low probability events like the threat of a fast-spreading deadly virus. We present the results of an ecological risk communication experiment conducted three weeks after the first case of Ebola was confirmed in the United States. The experiment investigates the influence of visual aids, numeracy, and emotions on judgments of diverse United States residents. It demonstrates how well-designed risk communication, numeracy, and emotions in synergy can exert powerful effects effects on informed decision making about personal and public health.

Chapter 5 illustrates how numeracy, emotions, and attitudes can affect decision making about health in another context: surrogate decision making.

In this chapter, we present the results of a decision making experiment conducted in the Netherlands simulating a decision about a sexually-transmitted virus like the Human Papilloma Virus (HPV). It shows how numeracy, emotions, and our own attitudes about risk can influence what decisions we would make as decision surrogates, in comparison to decisions we would make for ourselves.

Section II: Field Research

Chapter 6 reports the results of a field study conducted in Spain examining actual decisions of patients with cardiovascular disease (i.e., acute coronary syndrome). The study investigates the relationship between numeracy and decision delay in seeking care for symptoms. Results highlight the benefits of numeracy for efficient decision making when the stakes are high and time is of the essence.

Chapter 7 reports the results of a second field study in Spain examining actual patient outcomes. This second study investigated the effect of emotions, in particular the stable tendency to experience strong negative emotions but inhibit their expression, on health in patients with acute coronary syndrome. The results of this study suggest several biological and behavioral mechanisms through which emotions can impact health.

Chapter 8 examines healthy behavior. In this chapter, we present the results of a representative national cross-sectional survey in Spain. We examine the influence of perceptions of social support on adherence to screening for cardiovascular risk. The results demonstrate that social support can have profound effects on health, and that these effects are independent of and comparable to other traditional determinants of health.

Chapter 9 also examines actual health outcomes. In this chapter, we present the results of a meta-analysis of the effectiveness of programs designed to reduce incidence of sexually-transmitted infections among

United States adolescents. The results demonstrate the importance of providing comprehensive risk information and training in essential skills to make good decisions about health. In addition, we specify what type of information should be delivered in these programs to effectively reduce incidence of these infections.

Chapter 10 summarizes the results of an international qualitative study about risk communication regarding the Human Papilloma Virus vaccine. The findings demonstrate that transparent, trustworthy information about benefits and risks is an expected and desired part of health communication and decision making. Results also show that perceived risks and benefits are highly subjective, and that risk communication is often interpreted in context of people's expectations, and individual and social circumstances.

In the **Discussion**, I summarize the results of the research mentioned above and elaborate on its implications for theory, practice, and designing effective interventions.

References

Anderson, B. L., & Schulkin, J. (Eds.). (2014). *Numerical Reasoning in Judgments and Decision Making about Health*. Cambridge University Press.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior*, *31*(2), 143-164.

Barry, M. J., & Edgman-Levitan, S. (2012). Shared decision making—the pinnacle of patient-centered care. *New England Journal of Medicine*, *366*(9), 780-781.

Betsch, C., Böhm, R., Airhihenbuwa, C. O., Butler, R., Chapman, G. B., Haase, N., ... & Nurm, Ü. K. (2015). Improving medical decision making and health promotion through culture-sensitive health communication: an agenda for science and practice. Medical Decision Making. Advance online publication. doi: 10.1177/0272989X15600434

Borzekowski, D. L., Guan, Y., Smith, K. C., Erby, L. H., & Roter, D. L. (2013). The Angelina effect: immediate reach, grasp, and impact of going public. *Genetics in Medicine*, *16*(7), 516-521.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, *7*(1), 25-47.

Cokely, E.T., & Kelley, C.M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4,* 20-33.

Coulter, A., & Collins, A. (2011). *Making shared decision-making a reality: no decision about me, without me.* King's Fund.

DiClemente, R. J., Crosby, R. A., & Kegler, M. (Eds.). (2009). *Emerging theories in health promotion practice and research.* John Wiley & Sons.

Evans, D. G., Barwell, J., Eccles, D. M., Collins, A., Izatt, L., Jacobs, C., ... & Thomas, S. (2014). The Angelina Jolie effect: how high celebrity profile can have a major impact on provision of cancer related services. *Breast Cancer Research*, *16*(5), 442-448.

Evans, D. G., Wisely, J., Clancy, T., Lalloo, F., Wilson, M., Johnson, R., ... & Howell, A. (2015). Longer term effects of the Angelina Jolie effect: increased risk-reducing mastectomy rates in BRCA carriers and other high-risk women. *Breast Cancer Research*, 17, 143-144.

Ferrer, R. A., Klein, W. M. P., Lerner, J. S., Reyna, V., & Keltner, D. (2014). Emotions and health decision making: Extending the Appraisal Tendency Framework to improve health and health care. *Behavioral Economics and Public Health*, 1-25.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1-17.

Fishbein, M. (1979). A theory of reasoned action: Some applications and implications. *Nebraska Symposium on Motivation*, *27*, 65-116.

Fisher, J. D., & Fisher, W. A. (2002). The information-motivation-behavioral skills model. In R. J. DiClemente, R.A. Crosby, & M.C. Kegler, *EmergingTheories in Health Promotion Practice and Research: Strategies for Improving PublicHealth* (pp. 40-70). San Francisco, CA: Jossey-Bass.

Fowler, F. J., Levin, C. A., & Sepucha, K. R. (2011). Informing and involving patients to improve the quality of medical decisions. *Health Affairs*, *30*(4), 699-706.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied, 17,* 270-287*.*

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22(5),* 392-399.

Garcia-Retamero, R., & Cokely, E. T. (2014). The Influence of skills, message frame, and visual aids on Prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making, 27(2),* 179-189.

Garcia-Retamero, R., & Galesic, M. (Eds.). (2013). *Transparent communication of health risks: Overcoming cultural differences*. Springer Science & Business Media.

Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychology*, 33(8), 920-928.

Ghazal, S., Cokely, E.T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making*, *9,* 15-34.

Gigerenzer, G. (2012). Risk literacy. In J. Brockman (Ed.), *This will make you smarter: New scientific concepts to improve your thinking* (pp. 259-261). New York: Harper Perennial.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*(2), 53-96.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine*, *7*(7), e1000316.

Johnson, B. T., Carey, M. P., Marsh, K. L., Levin, K. D., & Scott-Sheldon, L. A. (2003). Interventions to reduce sexual risk for the human immunodeficiency virus in adolescents, 1985-2000: a research synthesis. *Archives of Pediatrics & Adolescent Medicine*, *157*(4), 381-388.

Jolie, A. (2013). My medical choice. *The New York Times*, *14*(05).

Kamenova, K., Reshef, A., & Caulfield, T. (2013). Angelina Jolie's faulty gene: newspaper coverage of a celebrity's preventive bilateral mastectomy in Canada, the United States, and the United Kingdom. *Genetics in Medicine*, *16*(7), 522-528.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. Davidson, K. Scherer & H. Goldsmith (Eds.), *Handbook of affective science* (pp. 619-642). New York: Oxford University Press.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267-286.

McKinley, S., Moser, D. K., & Dracup, K. (2000). Treatment-seeking behavior for acute myocardial infarction symptoms in North America and Australia. *Heart & Lung*, *29*(4), 237-247.

National Institute for Clinical Excellence. (2012). *Patient experience in adult NHS services: Improving the experience of care for people using adult NHS services.* NICE clinical guideline 138.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: theory and practice. *Annals of Behavioral Medicine*, *35*(3), 261-274.

Nieuwlaat, R., Wilczynski, N., Navarro, T., Hobson, N., Jeffery, R., Keepanasseril, A., ... & Sivaramalingam, B. (2014). Interventions for enhancing medication adherence. *Cochrane Database of Systematic Reviews, 11*, CD000011.

O'Connor, A. M., Bennett, C. L., Stacey, D., Barry, M., Col, N. F., Eden, K. B., ... & Llewellyn-Thomas, H. (2009). Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews*, *1,* CD001431.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making*, *25*, 390-401.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2015). Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *Journal of Experimental Psychology: Applied*, *21*, 178-194.

Petrova, D. G., Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: on the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making*, *27*(3), 191-199.

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education & Behavior*, *2(4),* 328-335.

Rimer, B. K., Briss, P. A., Zeller, P. K., Chan, E. C., & Woolf, S. H. (2004). Informed decision making: what is its role in cancer screening? *Cancer*, *101*(S5), 1214-1228.

Salzburg Global Seminar. (2011). The Salzburg Statement on Shared Decision Making. Retrieved on 03.02.2016 from https://www.lumc.nl/sub/2050/att/1405151019245212.pdf

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, *24*(2), 311-322.

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*(6), 322-325.

Trevena, L. J., Zikmund-Fisher, B. J., Edwards, A., Gaissmaier, W., Galesic, M., Han, P. K., ... & Ozanne, E. (2013). Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. *BMC Medical Informatics and Decision Making*, *13*(Suppl 2), S7.

van Steenkiste, B., van der Weijden, T., Timmermans, D., Vaes, J., Stoffers, J., & Grol, R. (2004). Patients' ideas, fears and expectations of their coronary risk: barriers for primary prevention. *Patient Education and Counseling*, *55*(2), 301-307.

Wegwarth, O., Schwartz, L. M., Woloshin, S., Gaissmaier, W., & Gigerenzer, G. (2012). Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Annals of Internal Medicine*, 156(5), 340-349.

SECTION I:

EXPERIMENTAL RESEARCH

CHAPTER 1:

A model of factors that shape informed decision making

The content of this chapter has been published as Petrova, D., Garcia-Retamero, R., & Cokely, E. (2015). Understanding Harms and Benefits of Cancer Screening: A Model of Factors that Shape Informed Decision Making. *Medical Decision Making*, *35*(7), 847-858. The supplemental data are available in this link: http://mdm.sagepub.com/content/35/7/847/suppl/DC1

Understanding Harms and Benefits of Cancer Screening: Model of Factors that Shape Informed Decision Making

Decisions about cancer screenings often involve consideration of complex and counterintuitive evidence. We investigated psychological factors that promote comprehension of benefits and harms associated with common cancer screenings and their influence on shared decision making. In Experiment 1, 256 men received information about PSA-based prostate cancer screening. In Experiment 2, 355 women received information about mammography-based breast cancer screening. In both studies, information about potential screening outcomes was provided in one of three formats: text, a fact box, or a visual aid (e.g., mortality with and without screening and rate of overdiagnosis). We modeled the interplay of comprehension, perceived risks and benefits, intention to participate in screening, and desire for shared decision making. Generally, visual aids were the most effective format increasing comprehension by up to 18%. Improved comprehension was associated with (1) better decision making (e.g., fewer intentions to participate in screening when it offered no benefit) and (2) more desire to share in decision making. However, comprehension of the evidence had a limited effect on emotions, risk perceptions, and decision making among those participants who felt that the consequences of cancer were extremely severe. Even when information is counterintuitive and requires integration of complex harms and benefits, user-friendly risk communications can facilitate comprehension, improve high-stakes decisions, and promote shared decision making. However, previous beliefs about the effectiveness of screening or strong fears about specific cancers may interfere with comprehension and informed decision making.

1. Introduction

Surveys show that the majority of American and European adults tend to be enthusiastic about cancer screening. Most people believe that cancer screenings are almost always beneficial and often grossly overestimate their benefits (Gigerenzer, Mata, & Frank, 2009; Hersch et al., 2013; Hoffman et al., 2010; Schwartz, Woloshin, Sox, Fischhoff, & Welch, 2000; Schwartz, Woloshin, Fowler, & Welch, 2004; Waller, Douglas, Whitaker, & Wardle, 2013). Many individuals are simply unaware that some screenings detect non-progressive cancers and may lead to substantial harms like unnecessary treatments (Hersch et al., 2013; Schwartz et al., 2000; Waller et al., 2013). When screenings show uncertain evidence of benefits, or when screenings are associated with substantial harms, experts recommend policies that promote informed and shared decision making (Esserman, Thompson, & Reid, 2013; Sheridan, Harris, & Woolf, 2004). Informed decision making requires that individuals understand the relevant benefits, risks, and limitations, consider their preferences, make a decision consistent with these preferences, and participate in decision making to the extent that they want (Rimer, Briss, Zeller, Chan, & Woolf, 2004).

Several obstacles to informed decision making about screening have been documented. For instance, physicians often fail to discuss potential harms of screening and fail to elicit their patients' preferences (Han et al., 2013; Hoffman et al., 2010; Wegwarth & Gigerenzer, 2011). The tension between benefits and harms can make decisions about screening cognitively taxing and psychologically difficult for all involved. Available evidence may be difficult to comprehend for some individuals (Arkes & Gaissmaier, 2012; Waller et al., 2013) and even when transparent "user-friendly" information is provided, some patients have difficulty integrating risks and benefits (Peters, Hibbard, Slovic, & Dieckmann, 2007; Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007; Peters, 2012). Compared to hard statistical evidence, anecdotes of early detection of cancer obtained through everyday sources (e.g., social networks) can be powerful motivators and may increase demand for screening even when early detection is not lifesaving (Fagerlin, Wang, & Ubel, 2005; Nowak & Parker, 2014). Somewhat unique to this context, the idea that a preventive behavior prescribed by experts can cause harm is surprising and counterintuitive for many individuals (Hersch et al., 2013; Waller et al., 2013), and may conflict with persuasive campaigns encouraging cancer screening without specifying the benefit or potential harms (Gigerenzer, 2014).

Given the increasing evidence of harms from screenings across a number cancers (e.g., overdiagnosis in breast, prostate, lung, and thyroid cancer screenings, (Esserman, Thompson, & Reid, 2013)) patients and doctors should understand the benefits and harms that will be a major part of many screening decisions. Unfortunately, although there are good general theoretical frameworks, there is not a substantial body of scientific literature investigating how people make evidence- and preference-based decisions in this context. How can we help people comprehend the controversial and sometimes counter-intuitive evidence about benefits and harms from some screenings? How does this evidence influence their desire to participate in screening and decision making about screening? To address these questions in an efficient manner, we conducted two risk communication experiments using representative and ecologically valid materials. In particular, we presented accurate information about common cancer screenings (e.g., prostate and breast cancer) to a diverse sample of individuals.

1.1. What factors influence comprehension?

Research suggests that physicians and patients have difficulty understanding screening statistics (Garcia-Retamero, Wicki, Cokely, & Hanson, 2014; Schwartz, Woloshin, Black, & Welch, 1997; Wegwarth, Schwartz, Woloshin, Gaissmaier, & Gigerenzer, 2012). In part, this difficulty may result from differences in skills and familiarity with numerical and probabilistic information, e.g., low numeracy and poor risk literacy (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Galesic M, 2010; Lipkus, Samsa, & Rimer, 2001; Peters, 2012; Reyna, Nelson, Han, & Dieckmann, 2009; Schwartz et al., 1997) (see www.RiskLiteracy.org). One factor that should influence comprehension of screening statistics is the format used to communicate benefits and harms (Trevena et al., 2013). Consider for example fact boxes that depict the most relevant information in a tabular format and have been effective in communicating benefits and harms in other contexts (Arkes & Gaissmaier, 2012; Schwartz, Woloshin, & Welch, 2009; Schwartz, Woloshin, & Welch, 2007; Woloshin, Schwartz, & Welch, 2004). Similarly, visual aids in the form of pictographs depict the number of affected and unaffected individuals using a matrix of icons (e.g., circles and faces) and facilitate comprehension in various contexts and populations, particularly among less numerate and more vulnerable populations (Gaissmaier et al., 2012; Galesic, Garcia-Retamero, & Gigerenzer, 2009; Garcia-Retamero & Galesic, 2010; Garcia-Retamero & Cokely, 2013; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012; Zikmund-Fisher et al., 2014; Zikmund-Fisher et al., 2008).

Although user-friendly formats may facilitate comprehension, the perceived severity of cancer (i.e., the degree to which people deem the consequences of a particular disease to be serious (Weinstein, 2000)) can reduce understanding. When the consequences of a decision are perceived to be serious or are affect-intensive (e.g., fear-inducing), decision makers tend to pay less attention to numerical, probabilistic information and rely more on heuristic-like processes that neglect the likelihood of specific events (Pachur, Hertwig, & Wolkewitz, 2014; Rottenstreich & Hsee, 2001). This suggests that individuals who perceive cancer as an extremely severe diagnosis may pay less attention to the evidence of screening effectiveness or else may give less weight to potential harms from screening.

1.2. How does comprehension relate to screening decisions?

Comprehension can increase patients' self-efficacy and perceived competence (Bandura, 1993). Comprehension can also help people realize the importance of their own preferences when careful, personal decision making is required (Hersch et al., 2013; Waller et al., 2013). Comprehension can promote understanding of the role of value judgments in addition to medical expertise (Kenny, Quine, Shiell, & Cameron, 1999), while lack of understanding can encourage patients to delegate decision making to others (Say, Murtagh, & Thomson, 2006).

Understanding that screening can cause serious harms, sometimes with minimal benefits, can dampen people's generally high enthusiasm for screening (Schwartz et al., 2004). When screening offers no benefits on average (e.g., prostate cancer screening (Ilic, Neuberger, Djulbegovic, & Dahm, 2013)), greater comprehension should be associated with intentions to avoid screening, in accord with expert recommendations (Moyer, 2012). Past research shows that men who reported being fully informed about advantages and disadvantages of prostate cancer screening tended to be less likely to undergo high-intensity screening (Han et al., 2013). Similarly, recent simulation modeling suggests that if patients learn the true likelihood that detection of early-stage breast cancer is lifesaving, screening rates may be reduced (Nowak & Parker, 2014). However, very weak relationships between comprehension and intentions to screen might suggest that other factors often have strong influences on screening decisions. These factors might include following a health professional's recommendation (Finney Rutten, Meissner, Breen, Vernon, & Rimer, 2005), strong fear of the disease, or discounting the presented information as a result of strong prior beliefs (Garcia-Retamero, Hoffrage, & Dieckmann, 2007; Garcia-Retamero, Müller, Catena, & Maldonado, 2009; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012).

1.3. Overview of experiments

We conducted two related experiments using a paid web panel group of diverse computer literate US residents (i.e., Amazon.com's Mechanical Turk)[1]. We studied the effects of information format and perceived severity on comprehension of benefits and harms from cancer screening. We modelled the influence of risk comprehension on participants' intentions to participate in screening and their desire to share decision making with one's physician. Given that age recommendations vary depending on the type of cancer screening, we invited adults of various ages to participate in the experiments. To be able to generalize the effects of comprehension beyond specific age groups, we controlled for age in our analyses. In Experiment 1, we communicated statistics about prostate cancer screening with PSA tests to men. According to many experts who recommend against early screening in the US, PSA screening does not reduce mortality and incurs a high risk of overdiagnosis (Ilic et al., 2013; Moyer, 2012). In Experiment 2, we tested the generalizability of our model and communicated statistics about breast cancer screening with mammography. According to experts, screening for breast cancer (a) offers modest but potentially life-saving benefits (Gøtzsche & Jørgensen, 2013) (b) results in a smaller rate of overdiagnosis compared to prostate cancer screening (Gøtzsche & Jørgensen, 2013; Ilic et al., 2013), and (c) may be better known to the public due to extensive media exposure and campaigns (Jolie, 2013; Thackeray, Burton, Giraud-Carrier, Rollins, & Draper, 2013).

2. Methods

2.1. Participants

---

[1] Mechanical Turk is a web panel of paid individuals that provides relatively diverse samples and relatively high-quality data roughly comparable to that provide by convenience and community or quota sampling (Paolacci & Chandler, 2014).

Experiment 1: Prostate cancer screening. Participants were 256 men (mean age=36, *SD*=13, range 18–70). Participants were Caucasian (81%), African American (7%), Hispanic (7%), Asian (4%), and other (1%). Participants had completed high school or less (12%), had at least some college education (27%), a 2-year college degree (13%), a 4-year college degree (35%), or a master's degree or higher (13%). Participants had various occupations including management (22%), unemployed (16%), working in service (11%), or sales/office (11%). Eighteen percent had been screened for prostate cancer at least once and 25% had talked to a health professional about prostate cancer screening. Three percent had been diagnosed with prostate cancer and 27% had a friend or a relative diagnosed with prostate cancer. Forty-six percent of participants reported not having heard of overdiagnosis before participating in the study, 40% knew that it existed but not much more, and 14% reported having extensive knowledge.[2]

Experiment 2: Breast cancer screening. Participants were 355 women (mean age=38, *SD*=14, range from 18 to 85) who were demographically similar to participants samples in Experiment 1. Thirty-six percent had been screened for breast cancer at least once and 44% had talked to a health professional about screening. Three percent had been diagnosed with breast cancer and 48% had a friend or a relative diagnosed with breast cancer. Thirty-one percent of participants reported never having heard of overdiagnosis before participating in the study, 49% knew that it existed but not much more, and 20% reported having extensive knowledge.

2.2. Design

Experiment 1 and 2 used the same design, measures, and procedure, and only differed in (a) gender of the sample of participants and (b) type of cancer and screening information presented. Participants were provided with

---

[2] The demographic and screening history questions for both experiments are available in online appendix C.

background information about prostate or breast cancer, respectively, adapted from the website of the US Centers for Disease Control and Prevention (available in online appendix A) and statistical information about benefits and harms from mammography or PSA test screening (Gøtzsche & Jørgensen, 2013; Ilic et al., 2013).[3] Participants in both experiments were randomly assigned to view the statistical information in one of three formats: text, fact box, or visual aids. Figure 1 shows the statistical information presented in the experiments and examples of the fact box and the visual aid. The full set of materials is available in online appendix B.

2.3. Dependent measures

Internal consistency coefficients and descriptive statistics for all measures used in Experiment 1 and 2 are presented in Table 1. Unless otherwise indicated, the items were developed specifically for this research.

Comprehension. Participants answered 8 questions designed specifically for this research (see Table S1 in the online supplement). Participants assessed comprehension of potential harm from screening (Q1-3) and of the statistical information presented about the degree of benefit and harm (Q4-8). The number of questions answered correctly showed good psychometric properties and was used as a measure of comprehension.

Emotional reactions. We assessed emotional responses to the communications of benefit and harm information with the Berlin Emotional Responses to Risk scale—an instrument currently in the final stages of validation for cross-cultural risk communication applications. It consists of items developed for health behavior research and has been used in published

---

[3] The statistics presented can vary depending on age group and other risk factors and the exact estimates are still under discussion (Barrat, 2015). However, our purpose was to convey possible extent of benefit and harm, so we presented averaged information. At the end of the study participants were debriefed and advised to consult a medical professional or a government website if they want to obtain more personalized risk estimates.

risk communication research since 2011 (Garcia-Retamero & Cokely, 2011; Garcia-Retamero & Cokely, 2014; Rothman, Martino, Bedell, Detweiler, & Salovey, 1999). The scale asks participants to indicate how they felt when reading the information about benefits and harms from screening. Specifically, on scales from 1 (not at all) to 7 (extremely) they indicated how assured, calm, cheerful, happy, hopeful, relaxed, relieved, anxious, afraid, discouraged, disturbed, sad, troubled, and worried they felt. The order of the adjectives was randomized. We averaged the scores across all negative adjectives as a measure of negative affect, and across all positive adjectives as a measure of positive affect.

Perceived severity of prostate/breast cancer. Participants also rated the seriousness of the consequences of having prostate/breast cancer on scales ranging from 1 (not at all serious) to 7 (extremely serious).

Perceived risk of prostate/breast cancer. Participants indicated on scales from 1 (not at all) to 7 (extremely) (1) how likely it was that they developed prostate/breast cancer in the next 11 years and (2) how worried they were that they would develop prostate/breast cancer in the following 11 years.

Perceived benefit of prostate/breast cancer screening. Participants indicated on scales from 1 (not at all) to 7 (extremely) (1) how effective prostate/breast cancer screening is in reducing the risk of dying from prostate/breast cancer, (2) how important it is to participate in prostate/breast cancer screening, and (3) how beneficial it is to participate in prostate/breast cancer screening.

Intentions. On scales from 1 (absolutely disagree) to 7 (absolutely agree) participants indicated to what extent they agreed with the following statements: (1) I intend to participate in prostate/breast cancer screening, and (2) I intend to look for information about prostate/breast cancer screening.

Shared decision making. We assessed participants’ estimated willingness to participate in decision making about prostate/breast cancer screening with an adjusted version of the Decision Making subscale of the Problem-Solving Decision-Making Scale (Deber, Kraetschmer, & Irvine, 1996). Participants indicated who should decide (1) how acceptable the risks and benefits of participating in prostate/breast cancer screening were, and (2) whether they would get screened or not. The answer options were (a) my doctor alone, (b) mostly my doctor, (c) my doctor and I equally, (d) mostly I, and (e) I alone. We computed a sum of the scores on two items and considered a score of 5-7 to indicate preference for shared, < 5 for delegated, and >7 for autonomous decision making.

Numeracy. We measured participants' numeracy with three items from Schwarz et al. (2007) and the adaptive version of the Berlin Numeracy Test (Cokely et al., 2012)—a test that is among the strongest predictors of one's ability to understand and make good decisions about risks (i.e., risk literacy; see RiskLiteracy.org for examples). Following Cokely et al. (2012), we used the sum of the participants' scores on both tests as our estimate of overall numeracy.

2.4. Procedure

The study was advertised as a study about "making decisions about health". Participants were eligible to take part in if they were US residents who were male (Experiment 1) or female (Experiment 2) and 18 years of age or older. Participants received 50 cents for their participation. They first read information about prostate/breast cancer and screening. Afterwards they viewed the statistical information about benefits and harms into one of the three formats: text, fact box, or visual aids. Then they answered the questions described above. The Ethics Committee of the University of Granada approved the methodology, and all participants consented to participation at the beginning of the study. There were no time constraints, but the entire survey took about 15 minutes to complete on average.

3. Results

We assessed effects of information format and perceived severity on comprehension in Experiment 1. We then examined correlations between comprehension, emotional reactions, perceptions of risk and benefit, intentions to participate in screening, and shared decision making. Finally, in a multiple regression framework we tested a series of mediation models to explain the relations between comprehension and decisions. We conducted similar analyses in Experiment 2 to test the generalizability of the model derived in Experiment 1. We controlled for demographic and other characteristics (e.g., age, education, numeracy) in the analyses. For the sake of brevity we report the results in the two experiments together. Descriptive statistics are shown in Table 1. The table also shows statistical comparisons between measures from the two experiments. The information provided in the two experiments differed on more than one dimension (gender, evidence, etc.). Investigating the effect of these dimensions was not theoretically central in the current research. We comment on these results whenever they are relevant to the main analyses.

3.1. Which information format facilitates comprehension?

On average, participants in both experiments correctly answered 62% [95% CI 60-64%] of the comprehension questions (65% [62- 69%] in Experiment 1 and 60% [58 to 62%] in Experiment 2). Participants perceived the consequences of cancer as moderately to extremely severe. Breast cancer was perceived as more severe than prostate cancer ($t=-6.47$, $p<.0001$, see Table 1). In both experiments, participants' perceptions showed negative skew with a median of 6 in Experiment 1 and a median of 7 in Experiment 2. We divided participants in two groups based on median split, such that 62% of participants in Experiment 1 and 60% in Experiment 2 were classified into the "extremely severe" vs. "moderately severe" group. We conducted analyses of variance (ANOVAs) with information format and perceived seriousness as independent variables and the number of correct comprehension questions as a dependent variable.[4]

Experiment 1. Information format had a significant effect on comprehension, *F*(2, 250)=4.06, *p*=.018, η_p^2=.03. Results show that visual aids increased comprehension (72% correct) compared to the textual message (64%, *p*=.050), and the fact box (61%, *p*=.003; see Figure 1SA of the online supplement). The fact box was not significantly different compared to the textual message (*p*=.314). Perceived severity had no effect on the number of correct comprehension questions (p>.1). Controlling for demographics did not influence the effect of format with *p*=.021, η_p^2=.03.

Experiment 2. Information format had no significant main effect, *F*(2, 349)=1.28, *p*=.281, η_p^2=.01. There was a marginally significant effect of perceived severity. Women who perceived breast cancer as extremely severe tended to have lower comprehension (*M*=4.67, *SD*=1.73) compared to women who perceived it as moderately severe (*M*=5.01, *SD*=1.80), *F*(1, 249)=3.09, *p*=.080, η_p^2=.01. There was also a marginally significant interaction between format and perceived severity, *F*(2, 349)=2.52, *p*=.082, η_p^2=.01 (see Figure 1SB in the online supplement). Consistent with results in Experiment 1, visual aids (67% correct) increased comprehension among people who perceived breast cancer as moderately severe compared to the other formats, *t*(143)=–1.62, $p_{(one\text{-}tailed)}$=.053 (fact box: 60% correct; text: 61% correct). However, for women who perceived breast cancer to be extremely severe, the textual message was the best format with 63% correct as compared to lower scores on both the fact box with 55% correct and the visual aid with 57%, *t*(208)= –2.34, *p*=.020. Controlling for demographics rendered the interaction between format and perceived severity significant with *p*=.046, η_p^2=.02.

---

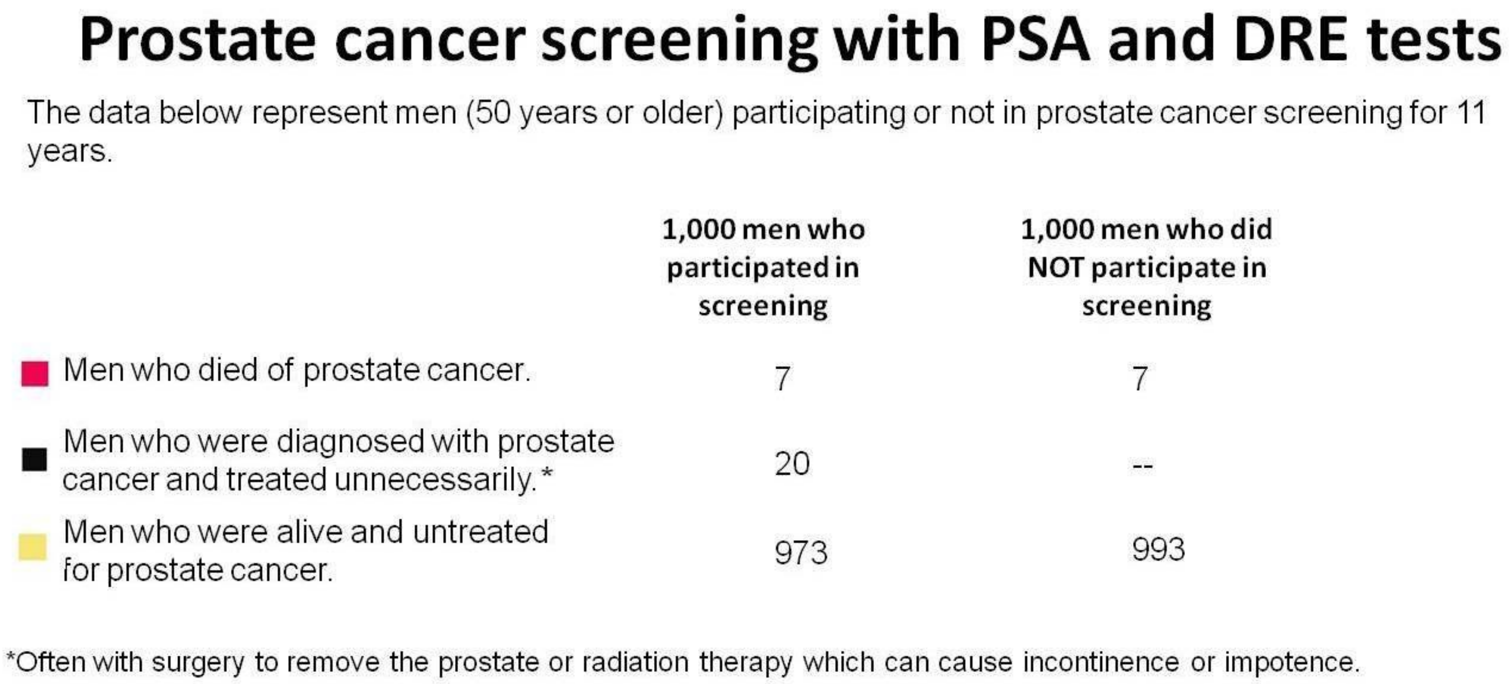
[4] Perceived severity was not a function of information format, *F*(352)=.009, *p*=.99.

Table 1. Descriptive statistics of the dependent measures, Cronbach's alpha for measures from combined items in Experiment 1 and 2, and statistical comparisons between the two experiments. Min.= minimum. Max.=maximum. SD=standard deviation.

	Experiment 1: Prostate Cancer Screening					Experiment 2: Breast Cancer Screening					T−test Results	
	Min.	*Max.*	*Mean*	*SD*	*Cronbach's alpha*	*Min.*	*Max.*	*Mean*	*SD*	*Cronbach's alpha*	*t*	*p*
Perceived severity of cancer	1	7	5.55	1.46	–	1	7	6.27	1.18	–	−6.47	<.001
Comprehension	0	8	5.23	2.04	.70	0	8	4.81	1.76	.57	2.71	.007
Positive affect	7	49	23.45	9.55	.88	4	45	21.97	8.82	.86	1.91	.056
Negative affect	6	49	22.63	10.35	.92	7	49	22.88	10.82	.92	−.28	.780
Perceived risk of cancer	2	14	5.54	2.94	.79	2	14	6.95	3.15	.82	−5.61	<.001
Perceived benefit of screening	3	21	11.24	5.69	.95	3	21	14.21	5.00	.93	−6.70	<.001
Intention to participate in screening	1	7	4.21	2.00	–	1	7	5.12	1.94	–	−5.62	<.001
Intention to look for more information	1	7	4.48	1.93	–	1	7	4.63	2.00	–	−.95	.340
Shared decision making	2	10	6.47	1.81	.84	2	10	6.81	1.53	.82	−2.55	.011
Numeracy	1	7	4.48	1.71	–	1	7	3.26	1.25	–	9.72	<.001

Figure 1. Statistical information about benefits and harms from screening communicated to participants. **A** is an example of the fact box used in Experiment 1 (based on Ilic et al., 2013). DRE=Digital rectal examination. **B** is an example of the visual aid used in Experiment 2 (based on Gøtzsche & Jørgensen, 2013). The fact box and the visual aid were designed after those provided by the Harding Center for Risk Literacy (http://www.harding−center.de).

A



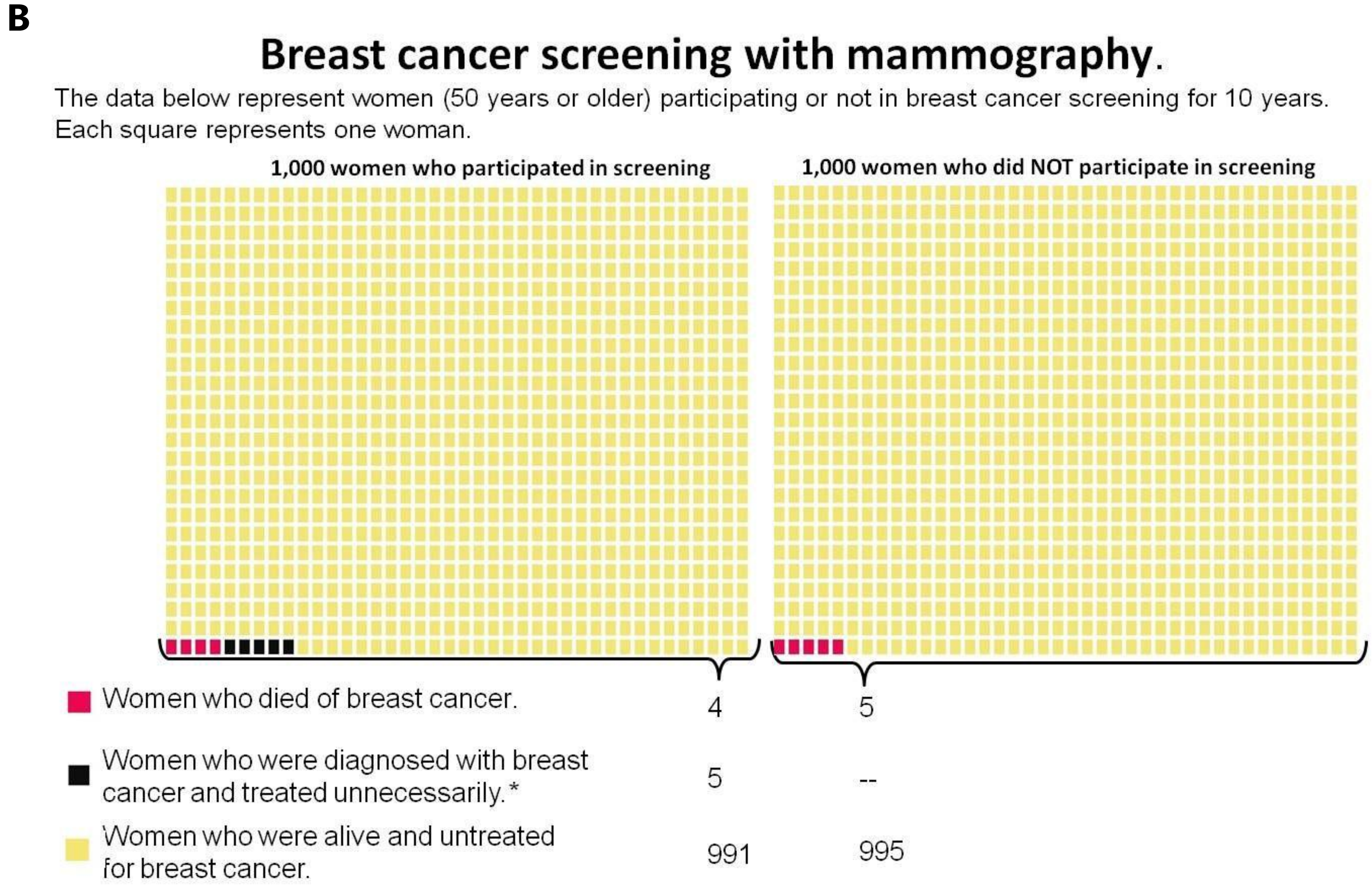
Prostate cancer screening with PSA and DRE tests

The data below represent men (50 years or older) participating or not in prostate cancer screening for 11 years.

	1,000 men who participated in screening	1,000 men who did NOT participate in screening
🟥 Men who died of prostate cancer.	7	7
⬛ Men who were diagnosed with prostate cancer and treated unnecessarily.*	20	--
🟨 Men who were alive and untreated for prostate cancer.	973	993

*Often with surgery to remove the prostate or radiation therapy which can cause incontinence or impotence.

Figure 1 continued.



*Often with mastectomy (partial or complete removal of the breast), radiation therapy or chemotherapy, which can cause fatigue and pain.

Figure 2. Intention to participate in screening (A) and desire for shared decision making (B) in Experiment 1 (Prostate cancer) and Experiment 2 (Breast cancer). For illustrative purposes, low comprehension was defined as ≤5 correct answers and high comprehension as >5 correct answers. In (A) a score of 4 marks the midpoint of scale (neither intends, nor does not intend to participate). In (B) a score between 5 and 7 marks a preference for shared decision making; score >7 indicates preference for active/autonomous decision making; score <5 indicates preference for delegated decision making. Error bars are ± 1 SEM.

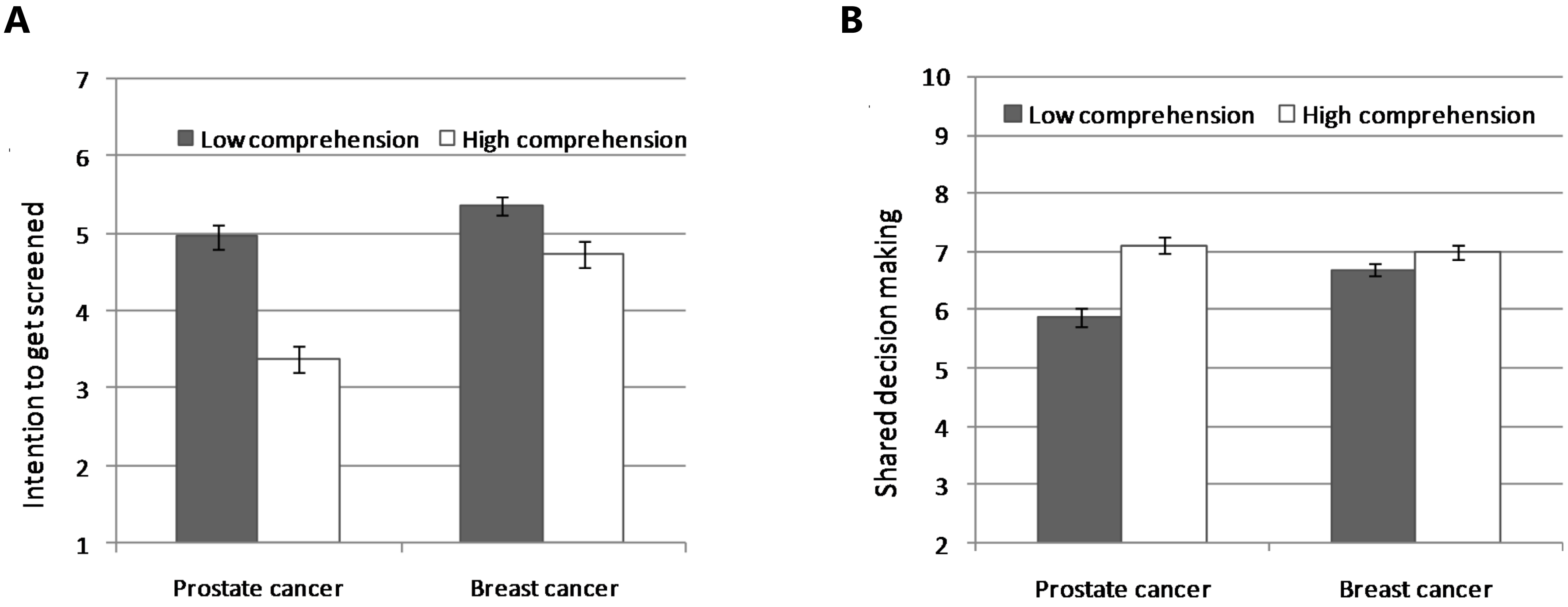


Figure 3. Mediation models. **I.** Experiment 1: Prostate cancer screening. **II.** Experiment 2: Breast cancer screening, **IIA**: Moderately severe group; **IIB**: Extremely severe group. Coefficients are Unstandardized B. The models control for age, education level, numeracy, information format, negative affect, and perceived severity of cancer (In Exp. 1). Dashed lines indicate non−significant paths (*p*≥.05).

I

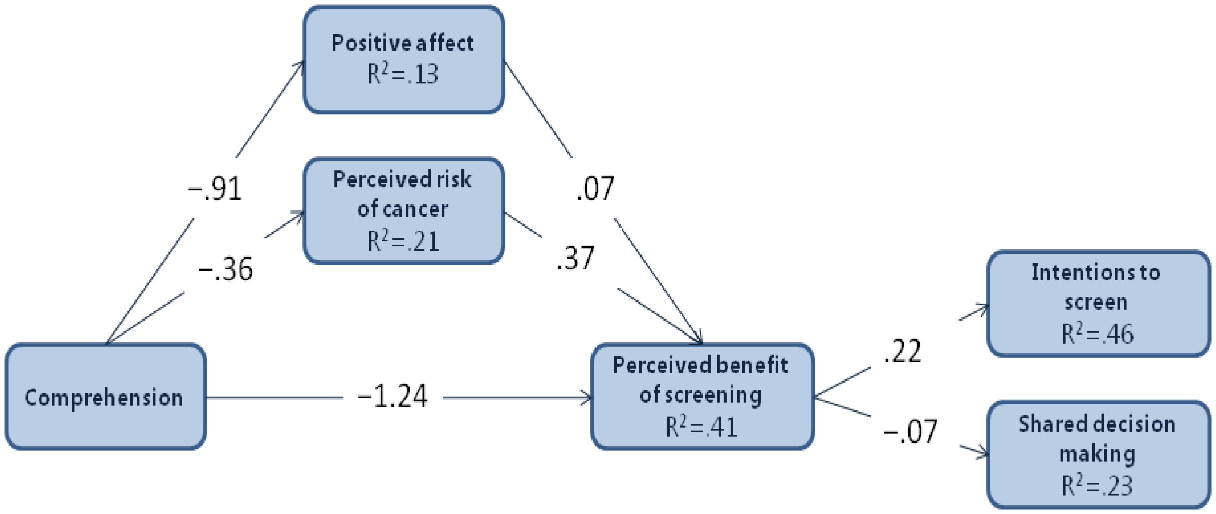
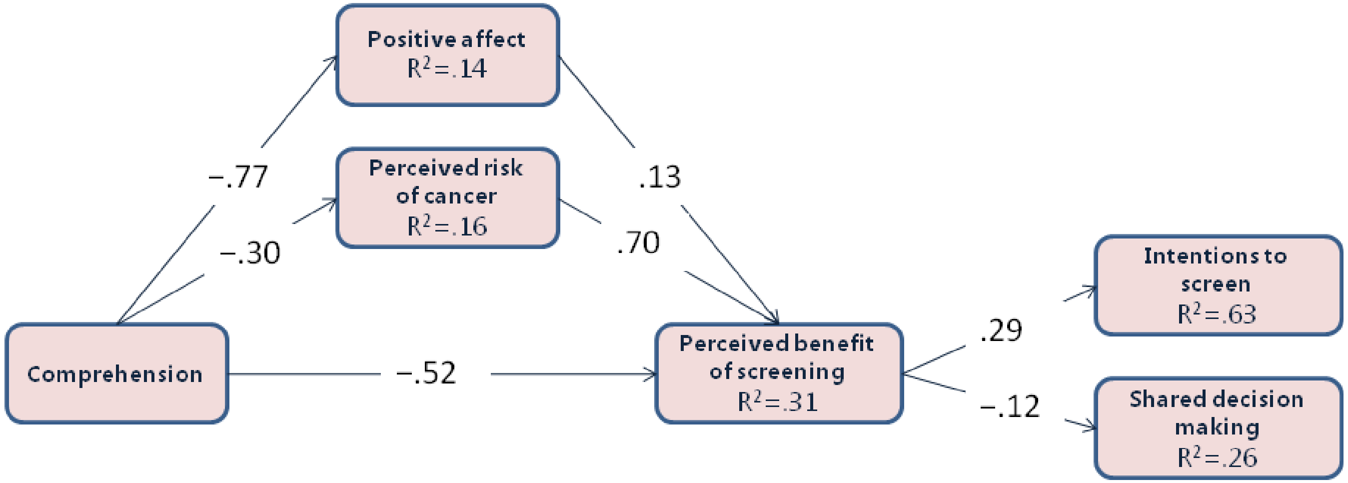
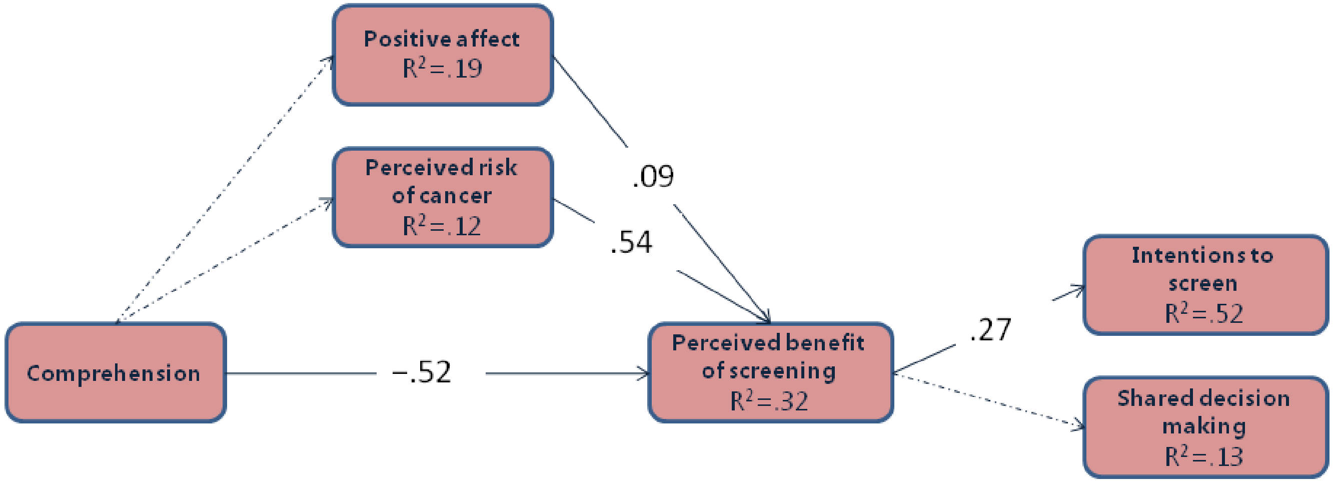


Figure 3 continued.

IIA



IIB



Finally, we checked whether having a relative or a friend diagnosed with prostate/breast cancer had an effect on comprehension or moderated any of the above mentioned effects. However, in both experiments there were no significant effects of this variable (*p*s>.05).

3.2. How is comprehension related to decisions?

Table 1 shows that on average participants in both experiments intended to participate in screening and preferred to share decision making with their physician. It is noteworthy that even when the screening statistics showed no benefits but showed substantial harm (Experiment 1), 44% of participants intended to participate in screening (indicated by a score>4) while 37% intended not to participate (score<4). When screening had both benefits and harms (Experiment 2), 66% intended to participate (score>4) while only 21% intended not to participate (score<4). Results also show that greater comprehension was related to less strong intentions to participate in screening and more desire to participate in decision making about screening (see Table S2 in the online supplement). Figure 2 shows that compared to participants who failed to understand the majority of the information, those who had good overall comprehension on average intended to forego screening in Experiment 1. In Experiment 2 high comprehension was also associated with less intention to get screened, although on average more participants intended to get screened. Figure 2 shows that despite the fact that comprehension was associated with an increase in the desire to participate in decision making, on average even participants with the highest comprehension still preferred to share decision making rather than be absolutely autonomous decision makers: People who understood the information the best still wanted to discuss options and consider the opinion of their physicians.

Greater comprehension was also related to lower positive affect, lower perceived risk of prostate cancer, and lower perceived benefit of screening (see Table S2). These variables were in turn related to the outcome variables (intentions to get screened and shared decision making), suggesting that they can be potential mediators of the relationship between comprehension and decisions.

It is also noteworthy that correlations between comprehension and outcome variables were consistently stronger in Experiment 1 than Experiment 2, showing that empirical evidence had a smaller impact on decisions in Experiment 2. In contrast, perceived severity was more strongly related to perceived benefit and intentions in Experiment 2 than in Experiment 1, showing that the perceived seriousness of cancer had a larger effect on decisions in Experiment 2.

3.3. Multifactorial process modeling

We used process modeling to assess how comprehension was related to decision making. Process modeling is an extension of mediation analysis that estimates direct and indirect effects in a multiple regression framework (56). We tested for indirect effects with a bias corrected bootstrap procedure using the PROCESS SPSS Macro (Hayes, 2008). Each model was based on 5000 bootstrap samples. We first analyzed data from Experiment 1, testing one model for each outcome (i.e., intention to get screened and shared decision making). Comprehension was included as a predictor and positive affect, perceived risk of cancer, and perceived benefit of screening were included as potential mediators (in this order). The choice of the candidate mediators was based on the presence of a significant correlation between the candidate mediator and the independent and outcome variables (see Table S2), and theoretical approaches emphasizing the influence of comprehension on decision making through emotions and perceived benefits and risks (Peters, 2012; Reyna et al., 2009). The order of mediators for the sequential effects was based on the sequence in which the measures were administered and the above-mentioned theoretical approaches. Each model tested for three simple indirect effects (i.e., through each single mediator) and four sequential indirect effects (i.e., through a sequence of two or all three mediators) operating simultaneously. In each model, we controlled for age, education level, numeracy, information format, negative affect, and perceived severity of cancer. We checked for mediation indicated by a significant total indirect effect. An indirect effect was considered significant if the 95% CI excluded 0.

Next, we sought to replicate the results from Experiment 1 with the data from Experiment 2. Because the perceived severity of breast cancer was associated with important differences in the results in Experiment 2, we estimated two separate models; one for the moderate and one for the extremely severe groups. The final results are displayed in Figure 3.

Experiment 1. In Experiment 1 there were significant total indirect effects of comprehension on intention to get screened, $M_{int}=-.33$ [95% CI −.43, −.25] and shared decision making $M_{sdm}=.12$ [.05, .21] (see Figure 3I). Participants who understood a larger proportion of the information perceived smaller risk of cancer, felt less relieved by the information about screening, and judged benefits of screening to be smaller. Consequently, they reported being less likely to get screened and reported more interest in shared decision making.

Experiment 2: Moderately severe group. Results for the group of women who judged breast cancer to be moderately severe were highly similar to the model results from Experiment 1. There were significant total indirect effects of comprehension on intention to get screened, $M_{int}=-.24$, [−.37, −.10] and shared decision making, $M_{sdm}=.10$ [.03, .19]. The same paths as in Experiment 1 emerged as significant (see Figure 3 IIA).

Experiment 2: Extremely severe group. The results for the group of women who judged the consequences of breast cancer to be extremely severe showed systematic deviations from previously estimated process models (see Figure 3 IIB). In this group, comprehension of the evidence had a small effect on decisions. There was a small total effect of comprehension on intentions to get screened, $M_{int}=-.16$ [−.28, −.06], and shared decision making, $M_{sdm}=.04$ [.004, .09]. Similar to the previous models, for the extremely severe group higher comprehension was associated with smaller perceived benefit of screening and less intention to get screened. However, comprehension was not related to positive affect or perceived risk of cancer. Regardless how well they understood the evidence, women who perceived breast cancer to be extremely severe reported being at high risk for breast cancer and reported feeling more assured and relieved upon reading the information about screening. These high risk perceptions and feelings of assurance were related to their increase in the perceived benefit of screening and their stronger intentions to screen.

4. Discussion

When benefits and harms were communicated following risk communication guidelines (Trevena et al., 2013) people understood a large proportion of the information correctly. Presenting the numerical information accompanied by a visual aid improved comprehension compared to alternative formats in participants who did not perceive the consequences of cancer as extremely severe (e.g., up to 18% relative improvement in Experiment 1). This result is consistent with previous research (Garcia-Retamero & Cokely, 2013; Zikmund-Fisher et al., 2014; Zikmund-Fisher et al., 2008; Zikmund-Fisher, Fagerlin, & Ubel, 2008) and suggests that simple visual aids can substantially improve risk comprehension even when risk communications involve complex, emotionally−charged, counter−intuitive evidence.

Surprisingly, the fact boxes used in this research did not reliably increase comprehension. Because information processing varies as a function of task complexity (Olshavsky, 1979; Payne, 1976),, there is reason to think that a structured tabular representation of information may primarily facilitate comprehension when the amount of information is larger. For example, in the current research we did not include other potentially relevant information, such as information about mortality from all causes or the proportion of false positive screening tests (Arkes & Gaissmaier, 2012). In cases when this information is highly–relevant and should be included, a fact box may be more beneficial.

In contrast to men's relatively moderate attitudes towards the severity of prostate cancer, a large proportion of women perceived breast cancer to be extremely severe. Our models indicated that the decisions of these women were less influenced by the available evidence. Women's feelings of assurance were not dampened by the evidence of harms from screening. Ironically, these feelings and perceptions were associated with more perceived benefit and more readiness to screen. Theoretically, these women could have given more weight to the benefits of screening than to the harms. Alternatively, decisions of these women may have been more influenced by pre–existing beliefs about the effectiveness of screening (Han et al., 2013) along with decision strategies (i.e., heuristics) stemming from the fear of disease (e.g., "prevention is always better"). Research shows that strong affective reactions related to the decision outcome (e.g., a potentially deadly disease) (Pachur et al., 2014; Rottenstreich & Hsee, 2001) or existing preconceptions based on previous information (Lewandowsky et al., 2012) can profoundly influence information processing, comprehension, and decision making when new information is presented. To illustrate, people find the idea of harms from screening surprising and counterintuitive (Hersch et al., 2013; Schwartz et al., 2004; Waller et al., 2013); at the same time many people report that screening is an obligation to one's family and society (Hersch et al., 2013; Schwartz et al., 2004; Waller et al., 2013). The presence of such strong previous beliefs and emotions may also explain why the subgroup of women who perceived breast cancer as extremely severe did not benefit from visual aids and even showed worse comprehension. Under conditions of more extreme emotional reactions to the cancer participants might have been less motivated to carefully study the information in an unfamiliar format, instead relying on established beliefs about effectiveness. These women may have also been distrustful of the information about harms, processing it more shallowly or discounting it as inconsistent with their beliefs (Lewandowsky et al., 2012).

In addition, people are likely to have little experience with serious harms resulting from preventive behaviors. This could explain why a large proportion of participants, and especially those who perceived breast cancer as extremely severe, were very enthusiastic about screening despite possible harms. To illustrate, people often know that drugs can cause side effects and this may reduce their willingness to take a drug (e.g., many women would choose not to take tamoxifen to reduce high breast cancer risk because of its side effects (Zikmund-Fisher et al., 2008). In contrast, personal accounts of people who were diagnosed with cancer after screening typically emphasize the idea that screening may have saved their life rather than that it may have caused them to undergo unnecessary surgery. Finally, the discrepant mammography recommendations issued by different official bodies in the US and the resulting controversy could have influenced some women's attitudes towards mammography and the provided information (Sharpe, Levin, Parker, & Rao, 2013). Future research should estimate the influences and processes by which non-evidence-based beliefs or anxieties create obstacles to informed decision making, and the means of overcoming these obstacles.

Several factors limit the generalizability of these findings. It should be noted that screening for cancer is usually recommended after the age of 50. Although approximately one third of our participants were among the age that was eligible for recommended screening, the participant samples also contained many younger participants for whom screening was not yet as relevant. In addition, Caucasian and highly educated individuals were overrepresented. Future research can replicate our study in probabilistic national samples and providing tailored information to participants consistent with their age and risk factors. Similarly, people who had some personal experience with prostate or breast cancer were overrepresented in our studies. While this might limit the generalizability of our findings, results showed that the benefits of comprehension were independent of age and education. While some preferences or experiences of our participants may not be fully representative of the intentions and preferences of people eligible for screening, our results suggest that the observed benefits of comprehension are likely to hold both within and beyond the target screening demographics. Nevertheless, future research should verify to what extent the effects of comprehension are similar among the populations underrepresented in our studies. Another potential limitation of the current set of studies is that perceived severity was measured after exposure to the information and so it may have been influenced by the way the information was presented. However, analyses showing a consistent lack of effect of format on perceived severity speak against this possibility, as does the fact that participants in all conditions received the same information about factors related to perceived severity (e.g., risk factors, symptoms, and treatment of the disease).

Previous research has shown that stable characteristics like age, gender, type of disease, and type of decision play a role in patients' preference for shared decision making (Arora & McHorney, 2000; Deber, Kraetschmer, Urowitz, & Sharpe, 2007; Say et al., 2006). The current findings add to this literature showing that a modifiable factor—i.e., comprehension (i.e., an essential component of informed decision making)—can promote willingness to participate in high−stakes value−sensitive decisions, independent of other influential factors (e.g., emotions and demographics). Broadly, the current research suggests that one's previous beliefs about the effectiveness of screening, emotions instilled by persuasive campaigns, and strong fears about specific cancers may interfere with shared and informed decision making. Results also suggested that user–friendly risk communications designed following expert guidelines (Trevena et al., 2013) may help attenuate the influence of these factors and more generally can improve high–stakes decisions while simultaneously promoting shared decision making. Just as one's comprehension of the harms and benefits of cancer screenings helps people make better choices and plans, perhaps understanding the need for informed decision making in one context will naturally translate into more participatory, informed decision making in other high–stakes domains.

5. References

Arkes, H. R., & Gaissmaier, W. (2012). Psychological research and the prostate-cancer screening controversy. *Psychological Science, 23*(6), 547-553.

Arora, N. K., & McHorney, C. A. (2000). Patient preferences for medical decision making: Who really wants to participate? *Medical Care, 38*(3), 335-341.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148.

Barratt, A. (2015). Overdiagnosis in mammography screening: A 45 year journey from shadowy idea to acknowledged reality. *BMJ, 350*, h867.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The berlin numeracy test. *Judgment and Decision Making, 7*, 25-47.

Deber, R. B., Kraetschmer, N., Irvine, J. (1996). What role do patients wish to play in treatment decision making? *Archives of Internal Medicine, 156*(13), 1414-1420.

Deber, R. B., Kraetschmer, N., Urowitz, S., & Sharpe, N. (2007). Do people want to be autonomous patients? Preferred roles in treatment decision-making in several patient populations. *Health Expectations, 10*(3), 248-258.

Esserman, L. J., Thompson, I. M., & Reid Jr., B. (2013). Overdiagnosis and overtreatment in cancer: An opportunity for improvement. *JAMA, 310*(8), 797-798.

Fagerlin, A., Wang, C., & Ubel, P. A. (2005). Reducing the influence of anecdotal reasoning on People's health care decisions: Is a picture worth a thousand statistics? *Medical Decision Making, 25*(4), 398-405.

Finney Rutten, L. J., Meissner, H. I., Breen, N., Vernon, S. W., & Rimer, B. K. (2005). Factors associated with men's use of prostate-specific antigen screening: Evidence from health information national trends survey. *Preventive Medicine, 40*(4), 461-468.

Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A., Broschinski, S., & Politi, M. C. (2012). Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. *Health Psychology, 31*(3), 286-296.

Galesic M, G. R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine, 170*(5), 462-468.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology, 28*(2), 210-216.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied, 17*(3), 270-287.

Garcia-Retamero, R., & Galesic, M. (2010). Who proficts from visual aids: Overcoming challenges in people's understanding of risks. *Social Science & Medicine, 70*(7), 1019-1025.

Garcia-Retamero, R., Hoffrage, U., & Dieckmann, A. (2007). When one cue is not enough: Combining fast and frugal heuristics with compound cue processing. *The Quarterly Journal of Experimental Psychology, 60*(9), 1197-1215.

Garcia-Retamero, R., Müller, S. M., Catena, A., & Maldonado, A. (2009). The power of causal beliefs and conflicting evidence on causal judgments and decision making. *Learning and Motivation, 40*(3), 284-297.

Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons’ preferred and actual roles in interactions with their patients. *Health Psychology, 33*(8), 920-928.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22*(5), 392-399.

Garcia-Retamero, R., & Cokely, E. T. (2014). The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making, 27*(2), 179-189.

Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women. *BMJ, 348*, g2636.

Gigerenzer, G., Mata, J., & Frank, R. (2009). Public knowledge of benefits of breast and prostate cancer screening in Europe. *Journal of the National Cancer Institute*, *101*, 1216-1220.

Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews,* CD001877.

Han, P. K. J., Kobrin, S., Breen, N., Joseph, D. A., Li, J., Frosch, D. L., & Klabunde, C. N. (2013). National evidence on the use of shared decision making in prostate-specific antigen screening. *The Annals of Family Medicine, 11*(4), 306-314.

Hayes, A. F. (2008). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hersch, J., Jansen, J., Barratt, A., Irwig, L., Houssami, N., Howard, K., . . . McCaffery, K. (2013). Women's views on overdiagnosis in breast cancer screening: A qualitative study. *BMJ, 346*, f158.

Hoffman, R. M., Lewis, C. L., Pignone, M. P., Couper, M. P., Barry, M. J., Elmore, J. G., . . . Zikmund-Fisher, B. J. (2010). Decision-making processes for breast, colorectal, and prostate cancer screening: The DECISIONS survey. *Medical Decision Making, 30*(5 Suppl), 53S-64S.

Ilic, D., Neuberger, M. M., Djulbegovic, M., & Dahm, P. (2013). Screening for prostate cancer. *Cochrane Database of Systematic Reviews,* CD004720.

Jolie, A. (2013). My medical choice. *The New York Times.* http://www.nytimes.com/2013/05/14/opinion/my-medical-choice.html. Accessed 13 Jan 2015.

Kenny, P., Quine, S., Shiell, A., & Cameron, S. (1999). Participation in treatment decision-making by women with early stage breast cancer. *Health Expectations, 2*(3), 159-168.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106-131.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*(1), 37-44.

Schwartz, L. M., Woloshin, S., Sox, H. C., Fischhoff, B., & Welch, H. G. (2000). US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: cross sectional survey. *BMJ, 320*(7250), 1635-1640.

Moyer, V. A. (2012). Screening for prostate cancer: US preventive services task force recommendation statement. *Annals of Internal Medicine, 157*(2), 120-134.

Nowak, S. A., & Parker, A. M. (2014). Social network effects of nonlifesaving early-stage breast cancer detection on mammography rates. *American Journal of Public Health, 104*(12), 2439-2444.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making, 25*(4), 390-401.

Olshavsky, R. W. (1979). Task complexity and contingent processing in decision making: A replication and extension. *Organizational Behavior and Human Performance, 24*(3), 300-316.

Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision, 1*(1), 64-78.

Paolacci, G., & Chandler, J. (2014). Inside the Turk: understanding mechanical turk as a participant pool. *Current Directions in Psychological Science, 23*(3), 184-188.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*(2), 366-387.

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs (Project Hope), 26*(3), 741-748.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31-35.

Peters, E., Dieckmann, N., Dixon, A., Hibbard, J. H., & Mertz, C. K. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review, 64*(2), 169-190.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943-973.

Rimer, B. K., Briss, P. A., Zeller, P. K., Chan, E. C., & Woolf, S. H. (2004). Informed decision making: What is its role in cancer screening? *Cancer, 101*(S5), 1214-1228.

Rothman, A. J., Martino, S. C., Bedell, B. T., Detweiler, J. B., & Salovey, P. (1999). The systematic influence of gain-and loss-framed messages on interest in and use of different types of health behavior. *Personality and Social Psychology Bulletin, 25*(11), 1355-1369.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science, 12*(3), 185-190.

Say, R., Murtagh, M., & Thomson, R. (2006). Patients' preference for involvement in medical decision making: A narrative review. *Patient Education and Counseling, 60*(2), 102-114.

Schwartz, L. M., Woloshin, S., Fowler Jr, F. J., & Welch, H. G. (2004). Enthusiasm for cancer screening in the United States. *JAMA*, *291*(1), 71-78.

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2009). Using a drug facts box to communicate drug benefits and harmstwo randomized trials. *Annals of Internal Medicine, 150*(8), 516-527.

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2007). The drug facts box: Providing consumers with simple tabular data on drug benefit and harm. *Medical Decision Making, 27*(5), 655-662.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*(11), 966-972.

Sharpe, R. E., Levin, D. C., Parker, L., & Rao, V. M. (2013). The effect of the controversial US preventive services task force recommendations on the use of screening mammography. *Journal of the American College of Radiology, 10*(1), 21-24.

Sheridan, S. L., Harris, R. P., & Woolf, S. H. (2004). Shared decision making about screening and chemoprevention: A suggested approach from the US preventive services task force. *American Journal of Preventive Medicine, 26*(1), 56-66.

Thackeray, R., Burton, S. H., Giraud-Carrier, C., Rollins, S., & Draper, C. R. (2013). Using twitter for breast cancer prevention: An analysis of breast cancer awareness month. *BMC Cancer, 13*(1), 508.

Trevena, L. J., Zikmund-Fisher, B. J., Edwards, A., Gaissmaier, W., Galesic, M., Han, P. K., . . . Lipkus, I. (2013). Presenting quantitative information about decision outcomes: A risk communication primer for patient decision aid developers. *BMC Medical Informatics and Decision Making, 13*(Suppl 2), S7.

Waller, J., Douglas, E., Whitaker, K. L., & Wardle, J. (2013). Women's responses to information about overdiagnosis in the UK breast cancer screening programme: A qualitative study. *BMJ Open, 3*(4), e002703.

Wegwarth, O., & Gigerenzer, G. (2011). "There is nothing to worry about": Gynecologists' counseling on mammography. *Patient Education and Counseling, 84*(2), 251-256.

Wegwarth, O., Schwartz, L. M., Woloshin, S., Gaissmaier, W., & Gigerenzer, G. (2012). Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Annals of Internal Medicine, 156*(5), 340-349.

Weinstein, N. D. (2000). Perceived probability, perceived severity, and health-protective behavior. *Health Psychology, 19*(1), 65-74.

Woloshin, S., Schwartz, L. M., & Welch, H. G. (2004). The value of benefit data in direct-to-consumer drug ads. *Health Affairs, 23*(3), 293-303.

Zikmund-Fisher, B. J., Witteman, H. O., Dickson, M., Fuhrel-Forbis, A., Kahn, V. C., Exe, N. L., . . . Fagerlin, A. (2014). Blocks, ovals, or people? icon type affects risk perceptions and recall of pictographs. *Medical Decision Making, 34*(4), 443-453.

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2008). Improving understanding of adjuvant therapy options by using simpler risk graphics. *Cancer, 113*(12), 3382-3390.

Zikmund-Fisher, B. J., Ubel, P. A., Smith, D. M., Derry, H. A., McClure, J. B., Stark, A., . . . Fagerlin, A. (2008). Communicating side effect risks in a tamoxifen prophylaxis decision aid: The debiasing influence of pictographs. *Patient Education and Counseling, 73*(2), 209-214.

CHAPTER 2:
Factors that influence complex decisions

The content of this chapter will be published as Petrova, D., Garcia-Retamero, R., Catena, A., & van der Pligt, J. (in press). To screen or not to screen: What factors influence complex screening decisions? *Journal of Experimental Psychology: Applied.* The appendix is found in this link: https://db.tt/oBrqZIYH.

Chapter 2

To screen or not to screen:
What factors influence complex screening decisions?

Contrary to people’s intuitions, many screenings have both benefits and harms (e.g., unnecessary treatments). Statistical information is often provided to ensure informed decision making. However, few theoretical models have addressed the role of comprehension of such information in screening decisions. In an experiment, we studied how cognitive skills, emotions, and a priori beliefs affect comprehension of the evidence of benefits and harms from screening and intentions to get screened. Young adults (N=347) received information about a disease for which a screening test was available and numerical information about the benefits and harms from screening. Results showed that comprehension and perceptions of benefits are central to decisions; however, lay perceptions of harms along the screening cascade require further study. Numeracy, science literacy, and emotions can promote informed decision making by facilitating comprehension of the evidence. At the same time emotions and beliefs resulting from persuasive campaigns can have strong effects on screening intentions beyond the available evidence. To apply to screening procedures where informed decision making is recommended, theoretical models of screening decisions need to include comprehension of benefits and harms, and account for how cognitive skills, emotions, and beliefs influence comprehension and decisions.

1. Introduction

Developments in medicine and health services have made an increasing number of screening tests available to individuals of various ages. Screening now starts in the womb (e.g., prenatal screening for Down's syndrome) and continues until much later in life (e.g., cancer screening). It is a good guess that our children will live in a world abundant with genetic screening tests. Screenings are done on asymptomatic people to find those at increased risk of having a disease or disorder, with the purpose to prolong and/or increase quality of life (Grimes & Schulz, 2002). As such, to be considered effective, screenings need to show benefit in randomized controlled trials (e.g., they should reduce mortality in people who go through screening compared to those who do not).

To the surprise of many, not all screenings show enough benefit to be recommended by authorities. For example, screening for prostate cancer with the prostate-specific antigen (PSA) test is considered not to be life-saving on average and it is not recommended in a number of countries (Ilic et al., 2013; Moyer, 2012). Even more counter-intuitively, some screenings can cause harms to individuals and by extension to economies. These harms have recently begun to receive more attention and have been categorized into four broader types: physical harms, psychological harms, financial strain, and opportunity costs (Harris et al., 2014). While people may be used to drug treatments causing adverse effects, they may not expect that preventive procedures can be harmful. For example, to the surprise of some women, screening for breast cancer with mammography results in many false positive tests and causes a proportion of women to undergo unnecessary cancer treatment, including mastectomy (i.e., the so called overdiagnosis bias; Gøtzsche & Jørgensen, 2013; Waller, Douglas, Whitaker, & Wardle, 2013).

Where difficult trade-offs between potential benefits and harms from screening need to be considered, experts recommend policies that promote informed decision making (Esserman, Thompson, Reid, 2013; Rimer, Briss, Zeller, Chan, & Woolf, 2004; Sheridan, Harris, & Woolf, 2004). This means that patients need to consider the relevant benefits, harms, risks, and limitations, and make a decision consistent with their preferences (Rimer et al., 2004). The information that should be communicated to patients frequently includes complex medical terms and probabilistic, numerical information about the evidence of benefits and harms. Comprehension of such information is potentially central to informed decision making.

However, there is no encompassing research framework that addresses the impact of various psychological factors on comprehension and decisions about screening when numerical information about both benefits and harms needs to be considered. For example, several health behavior models have been utilized in the context of cancer screening decisions—including the Health Belief Model (Rosenstock, 1974) and the Theory of Reasoned Action (Fishbein, 1979). These theories have identified important constructs that affect screening intentions and behavior. Some examples are perceived benefits and costs of screening, beliefs about screening, and perceived severity of the disease that screening might detect. However, these theoretical models did not consider the role of comprehension of health-relevant information. Other health behavior models like Social Cognitive Theory (Bandura, 2004) or the Information-Motivation-Behavioral Skills model (Fisher & Fisher, 2009) have addressed the role of health relevant knowledge in the prediction of health behavior. However, to the best of our knowledge, these models have been mostly based on and applied in the context of persuasion-based health promotion (i.e., encouraging a health behavior that is deemed desirable by experts). They have also not emphasized the role of numerical risk and benefit information which is common for decisions where informed decision making is recommended. In contrast, Fuzzy Trace Theory (Reyna, 2008), a model applied to medical decision making, gives a central role to comprehension. Fuzzy Trace Theory posits that decision makers rely on the gist of information (i.e., its bottom-line meaning) as opposed to verbatim details (i.e., the precise numbers). Research in physicians recently showed that gist knowledge was related to physicians' perceptions of benefits and harms from cancer screening, emphasizing the importance of comprehension in this context (Elstad et al., in press). However, while Fuzzy Trace Theory offers a comprehensive memory-based account of comprehension and how it affects decisions (Reyna, 2014), it does not make predictions about factors that may influence decisions without affecting gist or verbatim comprehension of the benefits and risks.

Other research showed how improved comprehension of benefits and harms can affect cancer screening intentions: Comprehension was related to more desire to participate in decision making about screening and better decisions (e.g., no intention to participate in a cancer screening program that offered no benefits; Petrova, Garcia-Retamero, & Cokely, 2015). Nevertheless, the effect of comprehension was limited among individuals who perceived cancer as extremely severe. This result suggests that beyond comprehension, emotional reactions to the prospect of a diagnosis and beliefs about screening can affect decision making, including the way in which individuals approach information about screening outcomes (Peters, Lipkus, & Diefenbach, 2006; Slovic, Peters, Finucane, & MacGregor, 2005). The psychology of modern screening decisions may benefit from a model that gives a central role to comprehension of the complex (numerical) information involved, and considers possible antecedents (e.g., emotional involvement, cognitive skills, attitudes) and consequences (e.g., screening intentions, behavior). The purpose of this research was twofold. Our first aim was investigating the influence of various factors—cognitive skills, emotions, and beliefs—on comprehension of screening statistics and screening intentions.

Our second aim was to start building a model of factors that influence complex screening decisions.

1.1. Cognitive skills

The necessity to comprehend complex numerical information in a medical context suggests that both numeracy and science literacy play an essential role in decision making. Numeracy refers to the ability to understand and operate with numerical and probabilistic concepts (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Galesic M, 2010; Lipkus, Samsa, & Rimer, 2001; Peters & Bjalkebring, 2015), a skill that might be essential when considering benefits and costs of screening participation. People with low numeracy overestimate risks and benefits of treatments and are less capable of adjusting their estimates from risk reduction information (Davids, Schapira, McAuliffe, & Nattinger, 2004; Garcia-Retamero & Galesic, 2010; Lipkus, Peters, Kimmick, Liotcheva, & Marcom, 2010; Reyna, Nelson, Han, & Dieckmann, 2009; Schwartz, Woloshin, Black, & Welch, 1997). People with low numeracy also search for less health-relevant information and prefer to leave decision making to their physician (Cokely et al., 2012; Galesic & Garcia-Retamero, 2011; Garcia-Retamero & Galesic, 2012; Garcia-Retamero, Wicki, Cokely, & Hanson, 2014). People with high numeracy, on the other hand, deliberate longer and make superior decisions (Ghazal, Cokely, & Garcia-Retamero, 2014).

While the predictive power of numeracy is well-established for various medical decisions, a substantial amount of variance in comprehension remains unexplained. The nature of the information and statistics about screening outcomes suggests that science literacy can also contribute to comprehension (Laugksch, 2000; National Science Foundation, 2014). In particular, understanding of how science generates and assesses evidence can help individuals evaluate the evidence of benefits and harms of screening (pp. 7-23, National Science Foundation, 2014). For example, when judging the effectiveness of screening, many people may consider all individuals alive in the screening group as "saved by screening"—even if there is a similar number of individuals alive in the non-screening group (Petrova et al., 2015). This result suggests that knowing the essence of experimental methods (e.g., a control group is required to establish the benefit of a treatment) can improve comprehension and adjust perceptions. Furthermore, good science literacy may encourage people to approach rather than avoid health-relevant numerical information. For example, people with adequate science literacy may have more experience with and greater liking for scientific information. This might increase deliberation (e.g., the amount of time that people spend reading relevant information) and facilitate comprehension. However, it is still unclear whether the influence of science literacy is independent from that of numeracy. To fill this gap, in this research we tested the following hypothesis:

Hypothesis 1(H1): Higher science literacy will be related to longer deliberation and better comprehension of benefits and harms of screening, and this effect will be independent from that of numeracy.

1.2. Emotions

Emotions can also affect screening intentions. Screening decisions are often emotionally laden because of the potential serious consequences (e.g., diagnosis of cancer; Zikmund-Fisher, Fagerlin, & Ubel, 2010). In the current research we tested two competing hypotheses about the effect of emotions on comprehension and decisions about screening. When the consequences of a decision are perceived to be emotionally powerful (e.g., fear-inducing), decision makers tend to pay less attention to numerical, probabilistic information and more often rely on heuristic-like processes that neglect the likelihood of specific events (Pachur, Hertwig, & Wolkewitz, 2014; Petrova, van der Pligt, & Garcia-Retamero, 2014; Rottenstreich & Hsee, 2001). This suggests that individuals who are more worried about a certain disease, or perceive the disease as more severe and unpleasant, may pay less attention to screening statistics and the actual evidence of benefits or harms. Instead, they would use heuristic-like strategies like “prevention is always good”. This implies that strong emotional reactions can be detrimental to comprehension and increase screening intentions, even when the evidence is not in favor of screening. Alternatively, people who are more worried about a certain disease may also be more motivated to understand the relevant evidence and make informed decisions. In other words, emotions can function as “spotlight” or “motivators”: they can increase people’s interest and motivation to process the relevant information (Peters et al., 2006), which can have a beneficial effect on comprehension and decisions.

Hypothesis 2a (H2a): Stronger emotional reactions will be related to worse comprehension and stronger screening intentions.

Hypothesis 2b (H2b): Stronger emotional reactions will be related to better comprehension and weaker screening intentions.

1.3. Beliefs

Beliefs about screening can also be related to comprehension and screening intentions. Many people can have strong positive beliefs about screening (e.g., many people think that screening is always a good choice; Schwartz, Woloshin, Fowler, &Welch, 2004; Waller, Osborne, & Wardle, 2015). These beliefs could rightfully stem from the perceived value of early prevention for saving people’s lives. These beliefs can be further enforced by exposure to persuasive campaigns encouraging regular screening. However, such campaigns rarely mention potential harms or specify the exact degree of benefit (e.g., the “pink ribbon” campaigns for breast cancer screening; Gigerenzer, 2014). As a result, people may be left with the impression that screening is useful by definition rather than a matter of choice based on a cost-benefit analysis. To illustrate, many US and European adults believe that cancer screenings are almost always beneficial and often grossly overestimate their benefits (Gigerenzer, Mata, & Frank, 2009; Hoffman et al., 2010;

Schwartz, Woloshin, Sox, Fischhoff, &Welch, 2000; Schwartz, et al., 2004; Waller et al., 2015). Similarly, many people perceive that screening is an obligation to one's family and society and perceive foregoing screening as irresponsible behavior (Hersch et al., 2013; Schwartz et al., 2004; Waller et al., 2015).

In our research we also tested two alternative hypotheses about the effect of beliefs on comprehension and screening intentions. In particular, non-evidence driven, prior beliefs can bias the processing of new information. For instance, prior beliefs might lead people to process new information about harms from screening more shallowly and/or discount the new, inconsistent information altogether (Garcia-Retamero, Müller, Catena, & Maldonado, 2009; Kunda, 1990; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). This implies that stronger positive a priori screening beliefs will reduce comprehension and, in turn, will increase screening intentions, even when the evidence is not in favor of screening. Alternatively, prior beliefs might not affect comprehension but only the way that people weigh potential benefits and harms from screening (Garcia-Retamero et al., 2009).That is, people with stronger positive beliefs might show similar levels of comprehension to those with less strong positive beliefs; however, they might focus on the benefits and ignore the harms when they make decisions about screening.

Hypothesis 3a (H3a): Stronger positive a priori beliefs will be related to lower comprehension and stronger screening intentions.

Hypothesis 3b (H3b): Stronger positive a priori beliefs will be related to stronger screening intentions, regardless of comprehension.

1.4. Theoretical model

We tested a model based on predictions derived from the Health Belief Model (Rosenstock, 1974), theories of numerical abilities, risk comprehension, and decision making (Reyna et al., 2009), and previous models of comprehension and informed decision making about screening (Petrova et al., 2015). People might differ in the amount of time that they spend deliberating on the statistical information that they receive. Based on previous research linking decision latency to superior decisions (Ghazal et al., 2014), we expected that shorter deliberation time would predict worse comprehension of screening outcomes, which in turn would predict more perceived benefits and fewer perceived harms from screening (Petrova et al., 2015). More perceived benefits and fewer perceived harms might be related to stronger intentions to get screened (Rosenstock, 1974). We further expected that participants would report stronger emotional reactions (e.g., fear, worry) for a disease described as more severe than for a disease described as rather neutral. These emotional reactions might be related to comprehension either negatively (H2a) or positively (H2b). We further tested whether individual differences in numeracy and science literacy would affect perceptions and intentions directly and indirectly (via comprehension). In particular, we expected numeracy and science literacy to be related to longer deliberations times and better comprehension (H1).We also tested whether a priori screening beliefs were related to stronger screening intentions indirectly (via comprehension; H3a) or directly (H3b).

To test this model under different conditions of screening effectiveness we manipulated the screening statistics. In one condition we presented participants with screening statistics showing that screening was effective at reducing mortality from the disease; in another condition screening was not effective. Hence, we investigated screening intentions when a normative decision option existed (i.e., not to participate in screening that is not effective) and under more ecological conditions where the extent of benefit needs to be weighed against the extent of possible harm (i.e., there is no "correct" decision option). We expected that the screening statistics manipulation would affect perceived benefits and intentions but not perceived harms (as the latter were held constant); we also accounted for a possible effect of the statistics manipulation on deliberation time and comprehension.

In sum, we investigated whether cognitive skills, emotions, and beliefs affect screening intentions by influencing comprehension of relevant evidence (i.e., *indirect effect on intentions via comprehension*). If these factors can affect screening intentions by impeding or facilitating comprehension of benefits and harms, results can help design interventions that promote informed decision making. In contrast, if these factors have a direct effect on screening intentions, results can help identify psychological processes influencing screening intentions beyond the statistical evidence.

2. Method

2.1. Participants

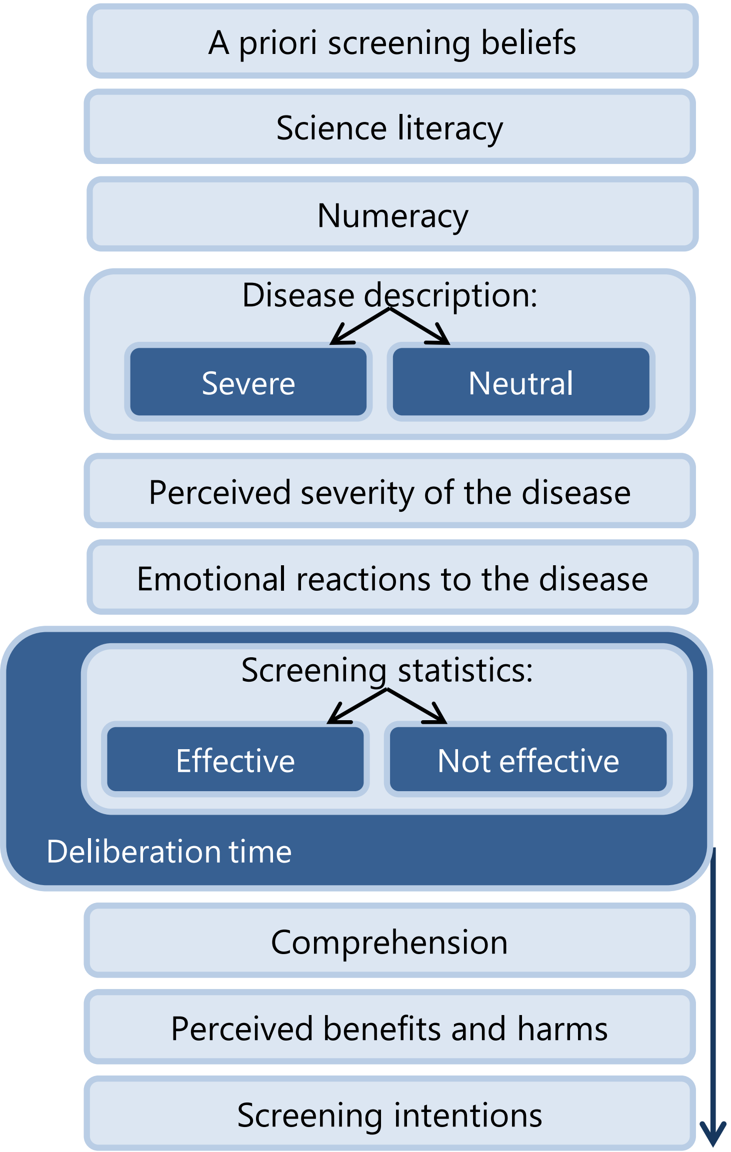
Participants were 347 first-year psychology students at the University of Amsterdam in the Netherlands (66% female, age *M*=20, *SD*=2.60). They completed the study at the beginning of their first academic semester (October 2014, i.e., before receiving any substantial training in psychology and scientific methods) as part of the University of Amsterdam test sessions for new students. Data collection was approved by the ethics committee at the Department of Psychology of the University of Amsterdam.

Stimuli and Design

Figure 1 shows a scheme of the design and the procedure of the experiment.

Disease description. Participants read a description of a hypothetical disease—Greene's disease—for which screening was available. Participants were told that the disease was discovered in Europe and affected around 5% of young adults. The disease was characterized by abnormal cells that were spread through the blood stream and could potentially cause death.

To investigate the effect of emotions, we manipulated the description of the disease. Half of the participants read a *severe* disease description, including extremely unpleasant symptoms of the disease and side-effects of the treatment. The rest of the participants received a relatively more *neutral* description of similar length that did not include information about the unpleasant effects. The descriptions were designed to evoke strong and weak negative emotional reactions, respectively (see the Appendix for the exact text of both conditions).



Screening statistics. All participants were told that there was a simple blood test available that could screen for Greene's disease. Participants read a short explanation about the potential benefits and harms of screening. They were told that the blood test was developed to detect the disease at an early stage, when the treatment had more chances of being effective. Participants were explained that the test could also detect some abnormal cells that do not threaten the person's life. However, doctors could not differentiate between dangerous and non-dangerous abnormal cells and people who test positive would get the treatment for Greene's disease even if they do not need it. Thus, some individuals would be treated unnecessarily.

All participants also read statistical information about the benefits and harms of screening for Greene's disease. The information was modeled after information available on government websites and decision aids about screenings with benefits and harms (see cdc.gov and harding-center.mpg.de). The statistics presented were fictitious but representative of the degree of harms and benefits from some cancer screenings (e.g., Gøtzsche & Jørgensen, 2013; Ilic, Neuberger, Djulbegovic, & Dahm, 2013). In particular, participants were told that scientists conducted a large-scale experiment to test the effectiveness of the blood test to save lives in 2,000 young people: One thousand were randomly assigned to participate in screening and another 1,000 were assigned not to participate in screening for 5 years.

Benefits. The information about the benefits of screening was manipulated between-subjects. Half of participants read information showing that *screening was effective* (i.e., saved lives). In particular, these participants were told: "Of 1000 young people who participated in regular screening for 5 years, *3 people died* of Greene's disease. Of 1000 young people who did not participate in regular screening for 5 years, *5 people died* of Greene's disease." The rest of the participants read information showing that *screening was not effective*. These participants were told: "Of 1000 young people who participated in regular screening for 5 years, *5 people died* of Greene's disease. Of 1000 young people who did not participate in regular screening for 5 years, *5 people died* of Greene's disease."

Harms. All participants received the same information about the degree of potential harm from screening. All participants were told: "From the 1,000 people who participated in screening, 6 people were treated with the laser treatment unnecessarily. From the 1,000 people who did not participate in screening, none were treated unnecessarily."

2.2. Measures

We measured a priori screening beliefs, cognitive skills (i.e., science literacy and numeracy), perceived severity of the disease (manipulation check), emotional reactions to the disease, deliberation (i.e., time that participants spent reading screening statistics), comprehension, perceived benefits and harms, and screening intentions.

A priori screening beliefs. On scales ranging from 1 (strongly disagree) to 7 (strongly agree), participants indicated to what extent they agreed with 6 statements that reflected positive attitudes towards screening (e.g., "Screening always has more advantages than disadvantages" and "Foregoing screening is irresponsible;" see Appendix for a full description). One item was excluded for low item-to-total correlation. The final score was the average of the remaining 5 items (Cronbach's alpha=.80) with *M*=4.46, *SD*=1.14.

Science literacy. It was assessed with three questions adapted from the US National Science Foundation survey (items 1-3, National Science Foundation, 2014) and two new items designed to complement the measure of the construct and increase reliability of the assessment (items 4-5; see Appendix for a full description). The final score was a sum of the number of correct answers on the 5 questions (Cronbach's alpha=.63) with a mean of 3.95 (*SD*=1.28).

Numeracy. It was assessed with 4 items from the Berlin Numeracy Test (Cokely et al., 2012; see www.RiskLiteracy.org). We computed the number of correct answers for each participant. The final score had a mean of 1.30 (*SD*=1.04).[5]

Perceived severity of the disease. Using scales from 1="not at all" to 7="most certainly", participants rated Greene's disease on 4 dimensions (severe, serious, unpleasant, and horrible). We computed the average score on the 4 items (Cronbach's alpha=.83). The final score had a mean of 5.25 (SD=1.11).

Emotional reactions to the disease. On scales from 1="not at all" to 7="most certainly" participants indicated to what extent they would feel afraid, worried, angry, hopeful, and calm if they had Greene's disease. The positive items were reverse-scored. We computed the average of the 5 items (Cronbach's alpha=.64) and the final score had a mean of 4.84 (SD=.83).

Deliberation time. We recorded how much time participants spent reading the screening statistics before they clicked "next" to see the comprehension questions. The comprehension questions were shown on a second page, and the screening statistics remained on the screen while participants answered them. However, participants were not informed that the statistics were going to remain on the screen while they answered the questions. Therefore, we used the time that participants spent reading the statistics on the first page as a proxy of deliberation. In order to correct for the positive skew typical for reaction time measures, the deliberation time measure was log-transformed for analysis.

---

[5] We conducted exploratory and confirmatory factor analysis to verify that the science literacy and numeracy items reflect two distinct constructs. Exploratory factor analysis returned a two-factor solution (Eigenvalues>1), explaining 40% of the variance, and consistent with each item loading strongly on its respective factor. A confirmatory factor analysis estimating two intercorrelated factors (r=.593) showed better model fit than a model estimating one single factor, as indicated by a drop in RMSEA and AIC coefficients (RMSEA .03 vs .05, AIC 93 vs. 105), confirming that the two scales capture two related but distinct constructs.

Comprehension. Participants answered 8 test-type questions that were adapted from previous research on screening with benefits and harms (see Appendix; see also Petrova et al., 2015). These items were designed to assess (1) comprehension of the idea that screening can have harms (items 1-3) and (2) comprehension of the most essential screening statistics (items 4-8).We computed the sum of correctly answered items. The items (Cronbach's alpha=.68) showed good discriminability with a mean of 4.19 (*SD*=2.11).

Perceived benefit of screening. On scales ranging from 1 (not at all) to 7 (very much) participants indicated (a) how effective screening with the blood test is, (b) how important, and (c) how beneficial it is to participate in screening for Greene's disease. The final score was an average of the three items (Cronbach's alpha=.84) with a mean of 4.05 (*SD*=1.30).

Perceived harm from screening. On scales ranging from 1 (not at all) to 7 (very much) participants indicated (a) how harmful and (b) how risky it is to participate in screening for Greene's disease. The final score was an average of the two items (Cronbach's alpha=.75) with a mean of 3.79 (*SD*=1.13).

Screening intentions. On scales ranging from 1 (strongly disagree) to 7 (strongly agree) participants indicated their agreement with the following statements: "I would regularly participate in screening for Greene's disease", "I would recommend to others to participate in screening for Greene's disease", "I would pay if necessary to participate in screening for Greene's disease". The final score was the average of the three items (Cronbach's alpha=.89) with a mean of 3.51 (*SD*=1.29).

In sum, we employed a 2 (disease description: severe or neutral) by 2 (screening statistics: effective or not effective) between-subjects design. Individual differences in a priori screening beliefs and cognitive skills (i.e., science literacy and numeracy) were considered as independent variables in the analyses. The main dependent variable was screening intentions, while emotional reactions to the disease, deliberation time, comprehension, and perceived benefit and harm were considered as potential mediators. Perceived severity of the disease was considered as a manipulation check.

2.3. Procedure

The experiment was introduced as a study about early detection of diseases through screening. Participants were provided with a short explanation of what screening is and were given several examples of prominent screening procedures (e.g., for cancer, high cholesterol, etc.). Participants then answered several questions about their beliefs about screening in general and filled in a science literacy test and a numeracy test. Then they received information about a disease, a screening procedure for this disease, and the benefits and harms from screening. Finally, participants answered a series of questions assessing emotional reactions, comprehension, perceived benefit of and harm from screening, and intentions to get screened.

2.4. Analysis

The goal of our analysis was to investigate whether and how the description of the disease, individual differences in science literacy and numeracy, a priori screening beliefs, and the effectiveness of screening influenced perceptions of screening and screening intentions. In particular, we investigated whether these variables influenced perceptions and intentions via promoting or impeding comprehension of evidence.

In the analyses, we first examined simple effects of the experimental manipulations (disease description and screening statistics) on the relevant dependent measures. In particular, we checked whether the description of the disease had the intended effect on the perceived severity of the disease, as well as on emotions, comprehension, perceptions of screening, and screening intentions. We also investigated whether the screening statistics manipulation had a direct effect on comprehension, perceptions of screening, and screening intentions. We then examined correlations between the three types of factors—cognitive skills, emotions, and beliefs—and deliberation time, comprehension, perceptions of screening, and screening intentions.

Finally, we conducted a path analysis in SPSS AMOS to test both direct and indirect effects in a multiple regression framework. We used General Least Squares estimation based on 500 bootstrap samples and calculated bias-corrected 95% confidence intervals (CI). To assess model fit we consulted the Root Mean Square of Approximation (RMSEA), the Chi-square test, and the Bayesian Information Criterion (BIC). In particular, RMSEA<.05 and a non-significant Chi-square test (p>.05) were used to assess the overall goodness of fit. The BIC was used for model selection because it takes into account both the statistical goodness of fit and the number of estimated parameters and imposes a strict penalty for increasing the number of parameters (Burnham & Anderson, 2002). In particular, a drop in BIC indicates an improvement in the model. We adopted a standard alpha level of .05 for all statistical decisions.

3. Results

3.1. Experimental manipulations

Disease description. Participants who read the severe disease description rated Greene's disease as more severe (*M*=5.41, *SD*=1.13) compared to participants who received the neutral description (*M*=5.15, *SD*=1.07), *t*(345)=−2.24, *p*=.026, Cohen's *d*=.24, showing that the manipulation was successful. Participants who read the severe description also reported stronger emotional reactions to the disease (*M*=4.96, *SD*=.99) compared to participants who read the neutral description (*M*=4.72, *SD*=.84), *t*(345)=−2.42, *p*=.016, Cohen's *d*=.26. The description had no direct effect on any of the other dependent measures, *t*s<1, *p*s>.5.

Screening statistics. Participants understood a similar proportion of the information regardless of whether screening was effective (*M*=4.14, *SD*=2.09) or not (*M*=4.24, *SD*=2.14), *t*(345)=.48, *p*=.633, Cohen's *d*=.05. Those who received information that screening saved lives perceived more benefit from screening (*M*=4.33, *SD*=1.12) compared to those who received information that screening saved no lives (*M*=3.78, *SD*=1.41), *t*(345)=−4.04, *p*<.0001, Cohen's *d*=.44. Those who received information that screening saved lives also had stronger intentions to get screened (*M*=3.66, *SD*=1.18 vs. *M*=3.36, *SD*=1.38), *t*(345)=−2.11, *p*=.035, Cohen's *d*=.23. There were no other significant effects, *t*s<2, *p*s>.2.Analyses of variance (ANOVA) showed that there were no significant interactions between screening statistics and the disease description on any of the mediator and outcome variables (*p*>.05).

3.2. Correlation analyses

Correlations between the individual difference measures and the dependent variables are shown in Table 1. Screening effectiveness might moderate some of the relationships shown in Table 1 (e.g., emotions may be more strongly related to intentions to get screened when screening is effective). Hence, before proceeding with the path analysis, we checked whether any of the above-mentioned significant relationships between the three factors (cognitive skills, emotions, and beliefs) and deliberation time, comprehension, perceived benefit of screening, and screening intentions depended on whether screening was effective or not. We centered the independent variables and computed the interaction terms between the variable screening effectiveness and the respective factor. Linear regression analyses showed no significant interactions for any of the above-mentioned relationships (all *p*s>.05), suggesting that they did not vary as a function of screening effectiveness.

3.3. Path analysis: Direct and indirect effects

Based on the correlation results (Table 1), we updated our conceptual model outlined in the introduction by allowing for covariance between (1) science literacy and numeracy and (2) science literacy and a priori screening beliefs (Model 1, a graphical illustration is found in the Appendix).

Table 1. Correlations between individual difference measures (cognitive skills, emotions, and beliefs) and dependent variables. *p≤.05

	Science literacy	Numeracy	Emotions	Deliberation	Comprehension	Perceived benefits	Perceived harms	Screening Intentions
Screening beliefs	.119*	-.023	.136*	.067	.004	.295*	.044	.228*
Science literacy		.353*	.249*	.475*	.387*	.028	.022	-.082
Numeracy			–.005	.261*	.300*	.035	-.053	-.063
Emotions				.253*	.208*	.124*	.153*	.095
Deliberation					.570*	-.081	.049	-.151*
Comprehension						-.166*	.014	-.239*
Perceived benefit							-.123*	.751*
Perceived harm								-.187*

Table 2. Indices for model comparison.

	RMSEA	BIC	df	Chi^2	p	Modification
Model 1	.07	296	27	68	<.001	-
Model 2	.05	233	39	76	<.001	Drop non-significant paths with p>.05 and abs(B)<.1.
Model 3	.04	221	38	58	.021	Estimate path from science literacy to emotional reactions.
Model 4	.03	218	37	49	.095	Estimate path from emotional reactions to perceived benefit.
Model 5	.02	214	36	39	.329	Estimate path from perceived harm to perceived benefit.

Note: df=model degrees of freedom. RMSEA=Root Mean Square of Approximation. BIC=Bayesian Information Criterion. Smaller values of BIC indicate a better model. Model 1=initial model. Non-significant paths dropped in Model 2 were: all effects on perceived harm, effects of a priori screening beliefs on deliberation, comprehension, and screening intentions; effects of screening statistics on comprehension and intentions; effects of science literacy on perceived benefits, and emotional reactions on comprehension.

g

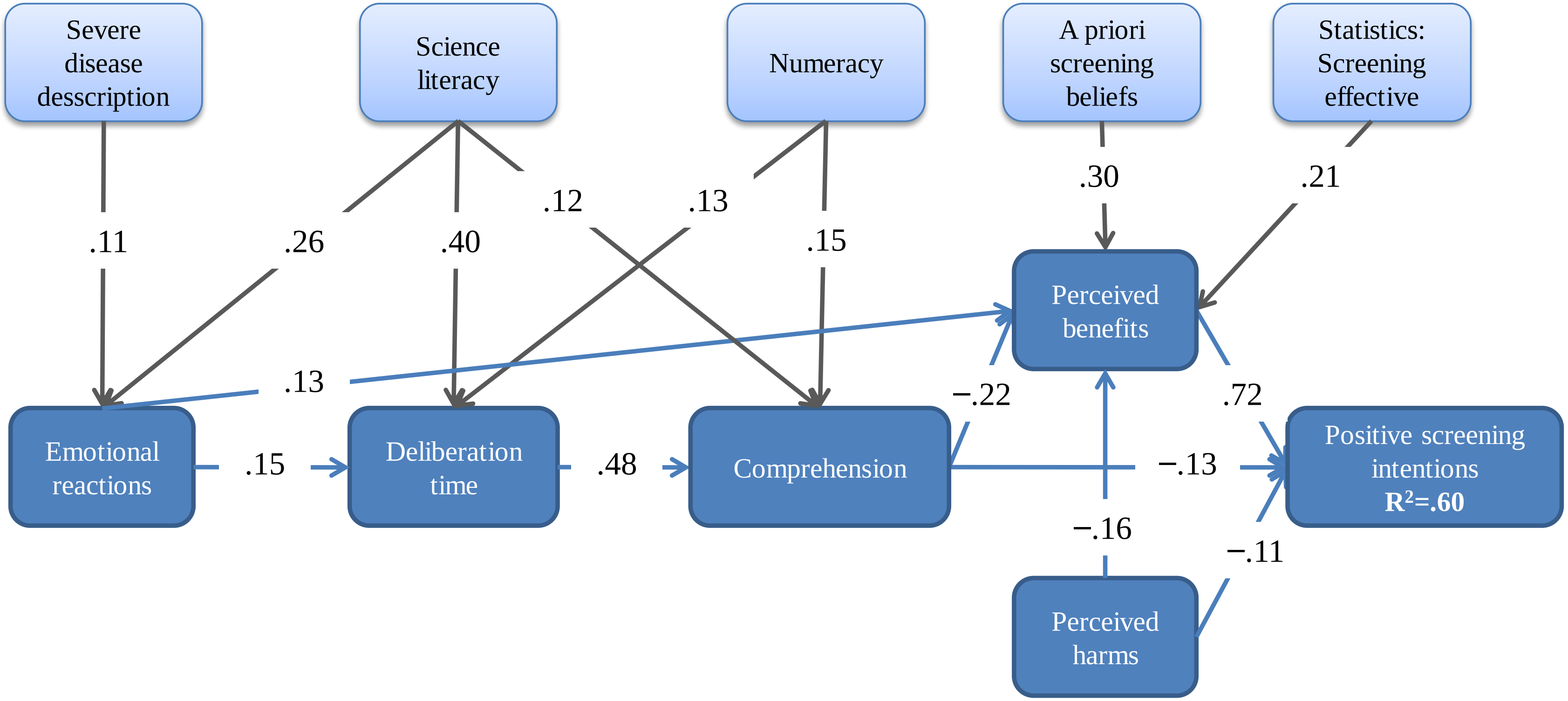

Severe disease desscription
Science literacy
Numeracy
A priori screening beliefs
Statistics: Screening effective
.11
.26
.40
.12
.13
.15
.30
.21
Perceived benefits
.13
−.22
.72
Emotional reactions
.15
Deliberation time
.48
Comprehension
−.13
Positive screening intentions
R²=.60
−.16
−.11
Perceived harms

Table 2 shows the fit indices for Model 1 and all subsequent model comparisons. Model 1 showed poor overall fit. In order to improve fit and give more power to the model, we started with fixing the non-significant paths to 0. In Model 2, all path weights from Model 1 that had an absolute value of the beta weight <.1 and *p*>.05 were dropped (i.e., fixed to 0), which improved model fit considerably. Next, we consulted modification indices. We considered freeing a parameter if the value of the index was >4, and the proposed relationship was theoretically meaningful. The largest modification index was for an effect of science literacy on emotional reactions, that was consistent with previous research documenting links between numeracy and emotional reactions (Peters et al., 2006; Petrova et al., 2014). Indeed, freeing this parameter improved model fit (see Model 3, Table 2). Next, we allowed for two changes that were consistent with the theoretical model of the "affect heuristic", stating that affective reactions towards an activity determine the perceived benefits and perceived risks associated with that activity (Slovic, Finucane, Peters, & MacGregor, 2004), and that these perceived benefits and risks are inversely related (Finucane, Alhakami, Slovic, & Johnson, 2000). In particular, we allowed for an effect from emotional reactions to perceived benefit (Model 4). Finally, we allowed for an effect of perceived harms on perceived benefits, which had a larger modification index (MI=11) than an effect of perceived benefits on perceived harms (MI=8) (Model 5). Both modifications improved model fit and the final Model 5 showed excellent overall goodness of fit. Figure 2 shows results from the final Model 5 that explained 60% of the variance in screening intentions.

Cognitive skills. Both science literacy, standardized effect (*SEF)*=.328[.225, .421], *p*=.004, and numeracy, *SEF*=.208[.111, .316], *p*=.004, had total positive effects on comprehension of screening statistics. Both science literacy, *SEF*=.12 [.015, .221], *p*=.035, and numeracy, *SEF*=.15 [.055, .251], *p*=.004, had significant direct effects on comprehension. In addition, science literacy was indirectly related to comprehension via both emotional engagement and deliberation time (see Figure 2). In contrast, numeracy was indirectly related to comprehension via deliberation time only (see Figure 2).

Emotions. There was a marginally significant indirect effect of the disease description on screening intentions, *SEF*=.008[95% CI −.001, .021], *p*=.071. Actually, the disease description had two distinct, opposite effects on intentions, which could explain the small total effect. The severe disease description was related to stronger emotional reactions (see Figure 2). On one hand, partially consistent with H2a (i.e., that emotions will be associated with worse comprehension and stronger screening intentions), these heightened emotional reactions were related to more perceived benefit of screening, *SEF*=.196, [.053, .372], *p*=.015, and to stronger screening intentions, *SEF*=.719, [.653, .777], *p*=.004, regardless of comprehension. On the other hand, consistent with H2b (i.e., that emotions will be related to better comprehension and weaker screening intentions), these heightened emotions were related to more time spent reading the information, which was related to better comprehension, *SEF*=.071, [.024, .133], *p*=.004, and weaker screening intentions, *SEF*=−.284, [−.381, −.191], *p*=.004.

Beliefs. Consistent with H3b (i.e., that beliefs will be related to stronger intentions regardless of comprehension), a priori screening beliefs were related to increased screening intentions via stronger perceptions of benefit, *SEF*=.212[.124, .293], *p*=.004, but had no effect on comprehension.

Finally, the statistical information that screening saved lives was related to stronger screening intentions via stronger perceptions of benefit, *SEF*=.151[.085, .224], *p*=.004(see Figure 2). Moreover, although perceived harm from screening was related to fewer perceived benefits and intentions to participate in screening, none of the variables in the model were related to perceived harms. Perceived harms were also a worse predictor of screening intentions than perceived benefits ($\text{Beta}_{\text{harms}}=-.11$ vs. $\text{Beta}_{\text{benefits}}=.72$, see Figure 2).

4. Discussion

Higher numeracy and science literacy independently predicted better comprehension, providing support for H1. A more severe disease description was related to more worry and fear from the disease. These heightened emotions had both beneficial and detrimental effects on comprehension and intentions to get screened. On one hand, stronger emotional reactions were related to more perceived benefits and stronger screening intentions; however, this effect was independent of comprehension, providing only partial support for H2a (i.e., that emotions would be related to worse comprehension and stronger screening intentions). On the other hand, consistent with H2b, stronger emotional reactions were related to longer deliberation time, better comprehension, and weaker screening intentions. Finally, our hypothesis that stronger positive a priory beliefs about screening would be related to worse comprehension (H3a) was not supported. Instead, positive a priori beliefs were associated with stronger intentions to get screened, regardless of the evidence of screening effectiveness and its comprehension, a result that is consistent with H3b. Finally, perceived benefits were more predictive of intentions than perceived harms.

4.1. Cognitive skills

Several studies have shown the beneficial effects of numeracy on risk comprehension in the context of screening (Davids et al., 2004; Galesic, Garcia-Retamero, & Gigerenzer, 2009; Lipkus et al., 2010; Reyna et al., 2009; Schwartz et al., 1997). However, to our knowledge this is the first study that shows that science literacy has a comparable, independent effect on comprehension. In this study, science literacy was operationalized roughly as knowledge of basic scientific methods used to derive conclusions (e.g., the belief that a control group is needed to establish treatment effectiveness and that correlation does not imply causation) (National Science Foundation, 2014). This knowledge is vital for evaluating benefits and harms from screening (e.g., knowing what comparison is important). It is also theoretically different from the ability to calculate risks and proportions captured by numeracy, although both concepts share some variance (Kahan et al., 2012; Schwartz, Woloshin, & Welch, 2005).

In addition, the path analysis showed that the two abilities influenced comprehension in slightly different ways. While both abilities had direct effects on comprehension, their indirect effects differed: numeracy was indirectly related to comprehension via deliberation time only, and science literacy was related to comprehension via both emotional reactions and deliberation time. In other words, the difference was that participants' numeracy was not related to their emotional reactions to the disease. This result may appear surprising at first sight, given that previous studies have related numeracy to affective reactions (Peters et al., 2006; Petrova et al., 2014). However, these studies have shown that numeracy is related to more precise feelings about numbers and higher emotional sensitivity to risks. That is, individuals with high numeracy have been often found to derive affective meaning from number comparisons (e.g., more differentiated feeling to 5% vs. 10% risk of disease, Peters et al., 2006; Petrova et al., 2014), something that was not made salient in the disease description. Instead, what was made salient were the characteristics and severity of the disease, which affected participants' emotions. It is then possible that higher science literacy contributed to deriving more meaning from the information about the nature of the fictitious disease and its treatment (e.g., because of better understanding of medical terms or knowledge similar diseases). This can explain why participants with higher science literacy perceived the disease as more frightening on average.

The indirect effect of science literacy on comprehension via deliberation time is in accordance with the hypothesis that participants with higher science literacy may be more likely to spend time processing the information (e.g., due to interest or familiarity with such type of information). In that sense, much like the way numeracy scales capture multiple numerical competencies (Peters & Bjalkebring, 2015), the science literacy scale may capture not only scientific knowledge per se but also people's need for cognition (i.e., their inclination towards effortful cognitive abilities; Cacioppo & Petty, 1982). Future research can investigate whether need for cognition and knowledge of scientific methods affect judgment and decision making independently.

Finally, although the science literacy items we used showed enough discriminability and good results despite a ceiling effect, the field could benefit from a more extensive validated and reliable measure of scientific reasoning. For example, the science literacy scale contains some numeracy-like items (see items 2 and 3 in the Appendix). In addition, the science literacy items were relatively easy and the numeracy items relatively difficult. This means that different performance on the two scales may not necessarily indicate differences in the underlying constructs but in the different difficulty levels. This limitation can be overcome in future research by using scales of comparable difficulty.

Note that the time spent on the page with statistical information is only a rough proxy for deliberation. Deliberation time may partially reflect one's perceived self-efficacy to comprehend the information, as well as one's experience and liking for this kind of information. The deliberation measure could also reflect an early selection metacognitive process, such that individuals who score high on cognitive abilities are more likely to process the information thoroughly and provide better monitoring during subsequent tasks (e.g., early selection vs. late correction; Cokely & Kelley, 2009, Ghazal et al, 2014). This might have helped these people find the correct answers once the questions were revealed. Alternatively, one could expect that people with higher cognitive abilities would process the information more efficiently and would actually spend less time on the page with information. Our results, however, do not support this alternative hypothesis. Nevertheless, while we think that the deliberation measure most likely reflects interest or more thorough strategic processing (see also Cokely& Kelley, 2009, Ghazal et al, 2014), it could also reflect conscientiousness or effort invested in the experiment in general. One way to deal with the limitations of this proxy measure in future research is to use process tracing methods like think-aloud protocols, which can offer a more direct insight into people’s deliberation.

4.2. Emotions

Evoking stronger negative emotional reactions towards the disease had two independent, opposite effects on screening intentions. On one hand, it was related to longer deliberation time and better comprehension, thereby decreasing intentions to get screened. One could say that in this case emotions had a *beneficial* effect on decision making, potentially facilitating informed decision making though increasing motivation and comprehension. This effect is also in line with the proposed function of affect as a “spotlight” or “motivator” (Peters et al., 2006). On the other hand, evoking stronger negative emotional reactions towards the disease was related to more perceived benefit of screening, thereby increasing intentions to get screened. Most importantly, this effect persisted regardless of comprehension and the effectiveness of screening, which were controlled for in the model. This means that participants who were more afraid of the disease perceived more benefit, even from screening that did not reduce chances of dying. In this case, this effect of emotions could be considered *detrimental* and leading to inferior, “non-informed” decision making. We should note that stronger emotional reactions were not related to lower comprehension, suggesting that they would not interfere with informed decision making by influencing how information is processed (e.g., paying less attention to numerical information and failing to answer the comprehension questions correctly). Rather, individuals who were more worried about the disease might have given more weight to the benefits of screening or followed a heuristic decision strategy (e.g., "always adopt a preventive behavior", "early detection is always good"). These results are in line with research on financial and medical decisions showing that when the consequences of an event are emotionally powerful, decision makers are less sensitive to the exact probability of its occurrence (in this case the probability of benefit or harm from screening) (Lejarraga, Pachur, Frey, & Hertwig, 2015; Pachur et al., 2014; Petrova et al., 2014; Rottenstreich & Hsee, 2001). Instead, they are more concerned about avoiding the focal negative outcome (in this case death from the disease). These results are also in line with a bulk of evidence suggesting that people process risk and benefit information not only cognitively but also emotionally, and that emotions are often even more influential in decisions about health risks (Loewenstein, Weber, Hsee, & Welch, 2001; Slovic, Finucane, Peters, & MacGregor, 2004; Zikmund-Fisher et al., 2010).

4.3. Beliefs

Much like the detrimental effect of emotions on intentions, positive a priori screening beliefs were related to more perceived benefit from screening and thus stronger intentions to get screened, regardless of comprehension and screening effectiveness. However, positive a priori screening beliefs had no effect on comprehension. Rather individuals who had more positive beliefs about screening (e.g., people who thought that screening was always beneficial or that foregoing screening was irresponsible) might have given more weight to the benefit of screening, even when there was statistically none. For instance, they could have given more weight to other potential benefits not captured by mortality reduction (e.g., the peace of mind in case of a negative screening result).

Our results showed that both strong a priori beliefs and emotional reactions are related to more perceived benefit of screening. This suggests that participants may have engaged in motivated reasoning (Kunda, 1990). Often when people are emotionally invested in a topic, they can trust evidence selectively in patterns that promote their goals and support their expectations; they can also discount or dismiss information that would cause them to experience dissonance or anxiety, especially when it is contrary to their beliefs (Lewandowsky et al., 2012; Kahan, 2012). This suggests that some participants (e.g., those who were very worried about the disease and those who believed screening was desirable) may have been distrustful of the counter-intuitive information about harms and lack of benefits. They could also have dismissed it altogether at the moment of decisions making because it did not validate their beliefs that screening is beneficial and did not relieve their fears from a severe disease. This potential mechanism may be even stronger in out-of-the-lab real screening decisions. If this is the case, previous research suggests that it may be difficult to re-adjust people's beliefs about the inherent goodness of screenings that have been reinforced by numerous screening campaigns (Lewandowsky et al., 2012).

National surveys show that many individuals share the positive a priori screening beliefs assessed in this study: they believe that screenings are almost always beneficial, overestimate their benefits, are unaware of potential harms, and perceive screening as an obligation to family and society (Gigerenzer et al., 2009; Hersch et al., 2013; Hoffman et al., 2010; Schwartz et al., 2000; Schwartz et al., 2004; Waller et al., 2015). The results of this study suggest that such beliefs can to an extent translate into greater willingness to get screened, even when the evidence is communicated in a simple, user-friendly way, and does not show clear benefits of screening. More broadly, persuasive screening campaigns encouraging screenings without specifying the extent of benefit or mentioning possible harms could contribute to "non-evidence-based" decision making, even when the evidence is provided. Future research should address this possibility with a more ecological approach.

4.4. Perceived benefits vs. perceived harms

Perceived benefits were central in the model, while perceived harms were less predictive of intentions and largely unrelated to cognitive skills, emotions, or beliefs. This suggests that in the context of screening or prevention in general "potential" motivation rather than "security" motivation may be the main driver behind decisions. In her seminal work in the 1980s, Lola Lopes proposed that risk-averse individuals are more motivated by security (e.g., they weigh the worst outcomes more heavily), while risk-seeking individuals are more motivated by potential (e.g., they weigh the best outcomes more heavily) (Lopes, 1987). The current results suggest that the importance of these motivations could be also context-dependent, such that when it comes to prevention (i.e., a gain context) people give more weight to potential benefits than to harms (see also Garcia-Retamero & Cokely, 2011). For instance, consistent with this context dependence, in the context of insurance decisions against a potential loss, negative emotions (e.g., fear) were more predictive of decisions than positive emotions (Petrova et al., 2014).

4.5. Limitations

Although we explained the concept of overtreatment to participants and referred to it as a harm resulting from screening, participants may not have considered overtreatment when answering the perceived harms questions. In other words, although participants may have considered the (unnecessary) treatment as very harmful, they may have failed to consider it as a consequence of screening (e.g., because given a positive screening result one can still choose not to undergo treatment). Consistent with this proposition, previous research suggests that some people see overtreatment as an issue of follow-up care rather than screening participation (Waller et al., 2013) and when naming harms from screening people more often focus on harms from the screening procedure itself rather than harms further along the screening cascade (Sutkowi-Hemstreet, et al., 2015). Perhaps if the screening procedure had been described as invasive, risky, or costly, perceived harm from screening would have shown stronger association with the other variables in the model. It is also possible that participants were simply very accepting of harms from overtreatment (i.e., did not consider it a deterrent to screening participation). Consistent with this, a probabilistic national survey in the United States found that 56% of respondents would want to be screened for pseudodisease. i.e., cancers that are slow growing or harmless and would not threaten the person's life in their lifetime (Schwartz et al., 2004). Finally, another possibility is that the made-up laser treatment, even in the high severity condition, was not perceived to be as invasive as the traditional cancer treatments like chemotherapy or surgery, which may be more likely to elicit decision-relevant perceptions of harm. Future research should investigate in more detail how lay perceptions of harms from screening are formed and in what way they are important for screening decisions.

In this study we measured comprehension with a set of items that assessed (1) understanding of the counterintuitive idea that screening can be harmful for some individuals, and (2) understanding of the benefits and harms of screening using numerical information. Most of the items in this study measured verbatim comprehension. Future research can extend our research by using items that measure gist comprehension (e.g., the difference between benefits and harms, Elstad et al., 2015). Beliefs and emotions may affect gist comprehension differently or to a greater extent than verbatim comprehension, and gist comprehension may be more predictive of decision than verbatim comprehension (Reyna, 2008; 2014).

Some limitations of this study were that the disease and screening were fictitious, the study population was homogenous, and was comprised of individuals who had little experience with health problems in general or screenings in particular. However, the evidence of benefits and harms from some real-world screenings is still under debate (Barratt, 2015) and is often not communicated to patients in campaigns or in clinical practice (Gigerenzer, 2014; Hoffman et al., 2010; Wegwarth & Gigerenzer, 2011), leaving some real-world decision makers just as naïve to the information as our participants. Moreover, the artificial context allowed us to experimentally manipulate the severity of the disease and the extent of benefit of screening. This would be difficult to achieve for a real disease like cancer, or a real screening program, with which participants have experience and where existing knowledge could influence responses. Also, it is important to keep in mind that the (statistical) information about screening presented to participants in this study may not be representative of all information that experts, patients, and other stakeholders may find relevant for decisions.

Some of the proposed hypotheses were not supported. In particular, we obtained no evidence suggesting that stronger emotional reactions (H2a) or stronger positive a priori screening beliefs (H3a) are related to worse comprehension. Differences between the hypothesized and the final model could reflect that the effects that were not observed had no bearing on screening intentions. However, as with any experimental research, it is possible that the observed model is specific to the scenario and participant population at hand. In order to build a comprehensive theory of complex screening decisions, our results should be replicated with more diverse populations in more ecological settings.

4.6. Implications for future research

The majority of studies investigating screening intentions and adherence have utilized theoretical frameworks that do not give a central role to comprehension of numerical risk and benefit information (e.g., the Health Belief Model, Rosenstock, 1974; and Theory of Reasoned Action, Fishbein, 1979). Instead, to address public health demands and experts' recommendations, most previous research on screening using these models has followed a persuasion-based tradition, looking to eliminate barriers to screening adherence (e.g., Austin, Ahmad, McNally, & Stewart, 2002; Curry & Emmons, 1994; Johnson, Mues, Mayne, & Kiblawi, 2008; Miller, Shoda, & Hurley, 1996). While such persuasion-based models may still apply to beneficial screenings with negligible harms, updated epidemiological evidence shows that many screenings can result in both benefits *and* serious harms (e.g., overdiagnosis in several cancer screenings, Esserman, Thompson, & Reid, 2013; fetal loss in prenatal screening, Lerman, Croyle, Tercyak, & Hamann, 2002; van den Berg, Timmermans, Leo, van Vugt, & van der Wal, 2006). For such screenings, policies promoting informed rather than persuasion-based decision making are recommended (Rimer et al., 2004; Sheridan et al., 2004; Woloshin, Schwartz, Black, & Kramer, 2012) and informed choice is a more justified indicator of success than uptake rates (Strech, 2014).

Achieving informed decision making about screening can be challenging because some medical concepts may be difficult to explain and the associated numerical evidence can be confusing. An informative example here is the case of harm from overdiagnosis, which we have simulated in this study. People consider information about overdiagnosis as complicated but important (Hersch et el., 2013; Waller et al., 2013). For example, research in women considering screening with mammography shows that screening intentions may depend heavily on the exact risk of overdiagnosis and that provision of numerical information about both screening benefits and harms from overdiagnosis increases informed choices (Hersch et al., 2013; 2015). While this highlights the importance of providing numerical information in general, it is especially important for future research to address how such information should be tailored so that each patient can benefit from it. The most effective format or the optimal amount of information provided may depend on the patient's specific information needs (Zikmund-Fischer, 2013) and numeracy (Garcia-Retamero & Cokely, 2013; Schwartz, 2011). For example, in a recent experiment a visual aid in the form of an icon array increased comprehension of benefits and harms from two common cancer screenings for many participants. However, it had an opposite effect among some participants who perceived the consequences of a cancer diagnosis as extremely severe (Petrova et al., 2015).

Given that comprehension of benefits and harms is central to informed decision making, we need updated theoretical models that apply to screening procedures where informed decision making is recommended. These should give a more central role to comprehension and predict how it influences decisions (e.g., Fuzzy Trace Theory, Reyna, 2008, 2014) and how it is influenced by relevant person- and situation-based factors. This study identified some factors whose effects may generalize to other decision making contexts where numerical information is involved (e.g., the role of emotions and science literacy). It also identified some issues that may be unique to complex screening decisions. One example is the non-evidence based influence of a priori beliefs about the goodness of screening on intentions. Another one is the unclear, potentially complex antecedents of perceived harm from screening that require further study in a more naturalistic setting.

5. References

Austin, L. T., Ahmad, F., McNally, M., & Stewart, D. E. (2002). Breast and cervical cancer screening in Hispanic women: A literature review using the health belief model. *Women's Health Issues, 12*(3), 122-128.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior*, *31*(2), 143-164.

Barratt, A. (2015). Overdiagnosis in mammography screening: A 45 year journey from shadowy idea to acknowledged reality. *BMJ (Clinical Research Ed.), 350*, h867.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach.* Springer.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116-131.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*, 25-47.

Cokely, E.T., & Kelley, C.M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*, 20-33.

Curry, S. J., & Emmons, K. M. (1994). Theoretical models for predicting and improving compliance with breast cancer screening. *Annals of Behavioral Medicine*, *16*(4), 302-316.

Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine, 19*(4), 310-315.

Elstad, E. A., Sutkowi-Hemstreet, A., Sheridan, S. L., Vu, M., Harris, R., Reyna, V. F., ... & Brewer, N. T. (2015). Clinicians' perceptions of the benefits and harms of prostate and colorectal cancer screening. *Medical Decision Making*, Epub ahead of print. DOI: 10.1177/0272989X15569780

Esserman, L. J., Thompson, I. M., Jr,&Reid, B. (2013). Overdiagnosis and overtreatment in cancer: An opportunity for improvement. *JAMA, 310*(8), 797-798.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1-17.

Fishbein, M. (1979). A theory of reasoned action: Some applications and implications. *Nebraska Symposium on Motivation*, *27*, 65-116.

Fisher, J. D., & Fisher, W. A. (2002). The information-motivation-behavioral skills model. In R. J. DiClemente, R.A. Crosby, & M.C. Kegler, *EmergingTheories in Health Promotion Practice and Research: Strategies for Improving PublicHealth* (pp. 40-70). San Francisco, CA: Jossey-Bass.

Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine, 170*(5), 462-468.

Galesic, M., & Garcia-Retamero, R. (2011). Do low-numeracy people avoid shared decision making? *Health Psychology, 30*(3), 336-341.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology, 28*(2), 210-216.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied*, *17*, 270-287.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science*, *22*(5), 392-399.

Garcia-Retamero, R., & Galesic, M. (2012). Doc, what would you do if you were me? on self–other discrepancies in medical decision making. *Journal of Experimental Psychology: Applied, 18*(1), 38-51.

Garcia-Retamero, R., & Galesic, M. (2010). How to reduce the effect of framing on messages about health. *Journal of General Internal Medicine, 25*(12), 1323-1329.

Garcia-Retamero, R., Müller, S. M., Catena, A., & Maldonado, A. (2009). The power of causal beliefs and conflicting evidence on causal judgments and decision making. *Learning and Motivation, 40*(3), 284-297.

Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychology, 33*(8), 920-928.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women. *BMJ: British Medical Journal, 348,* g263.

Gigerenzer, G., Mata, J., & Frank, R. (2009). Public knowledge of benefits of breast and prostate cancer screening in Europe. *Journal of the National Cancer Institute*, *101*, 1216-1220.

Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews,6*, CD001877.

Grimes, D. A., & Schulz, K. F. (2002). Uses and abuses of screening tests. *The Lancet, 359*(9309), 881-884.

Harris, R. P., Sheridan, S. L., Lewis, C. L., Barclay, C., Vu, M. B., Kistler, C. E., et al. (2014). The harms of screening: A proposed taxonomy and application to lung cancer screening. *JAMA Internal Medicine, 174*(2), 281-286.

Hersch, J., Jansen, J., Barratt, A., Irwig, L., Houssami, N., Howard, K., et al. (2013). Women's views on overdiagnosis in breast cancer screening: A qualitative study. *BMJ, 346*, f158.

Hersch, J., Barratt, A., Jansen, J., Irwig, L., McGeechan, K., Jacklyn, G., ... & McCaffery, K. (2015). Use of a decision aid including information on overdetection to support informed choice about breast cancer screening: a randomised controlled trial. *The Lancet*, *385*(9978), 1642-1652.

Hibbard, J. H., Peters, E., Dixon, A., & Tusler, M. (2007). Consumer competencies and the use of comparative quality information: It isn't just about literacy.*Medical Care Research and Review, 64*(4), 379-394.

Hoffman, R. M., Lewis, C. L., Pignone, M. P., Couper, M. P., Barry, M. J., Elmore, J. G., et al. (2010). Decision-making processes for breast, colorectal, and prostate cancer screening: The DECISIONS survey. *Medical Decision Making, 30*(5 Suppl), 53S-64S.

Ilic, D., Neuberger, M. M., Djulbegovic, M., & Dahm, P. (2013). Screening for prostate cancer. *Cochrane Database of Systematic Reviews,1*, CD004720.

Johnson, C. E., Mues, K. E., Mayne, S. L., & Kiblawi, A. N. (2008). Cervical cancer screening among immigrants and ethnic minorities: A systematic review using the Health Belief Model. *Journal of Lower Genital Tract Disease, 12*(3), 232-241.

Kahan, D. M. (2012). Ideology, motivated reasoning, and cognitive reflection: an experimental study. *Judgment and Decision Making*, *8*, 407-24.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732-735.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin,108*(3), 480-498.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing.*Psychological Science in the Public Interest*, *13*(3), 106-131.

Laugksch, R. C. (2000). Scientific literacy: A conceptual overview. *Science Education, 84*(1), 71-94.

Lejarraga, T., Pachur, T., Frey, R., & Hertwig, R. (2015). Decisions from experience: From monetary to medical gambles. *Journal of Behavioral Decision Making.*Online first: *DOI: 10.1002/bdm.1877*

Lerman, C., Croyle, R. T., Tercyak, K. P., & Hamann, H. (2002). Genetic testing: Psychological aspects and implications. *Journal of Consulting and Clinical Psychology, 70*(3), 784.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106-131.

Lipkus, I. M., Peters, E., Kimmick, G., Liotcheva, V., & Marcom, P. (2010). Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Medical Decision Making, 30*(4), 464-473.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*(1), 37-44.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267-286.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology, 20*, 255-295.

Miller, S. M., Shoda, Y., & Hurley, K. (1996). Applying cognitive-social theory to health-protective behavior: Breast self-examination in cancer screening. *Psychological Bulletin, 119*(1), 70-94.

Moyer, V. A. (2012). Screening for prostate cancer: US preventive services task force recommendation statement. *Annals of Internal Medicine, 157*(2), 120-134.

National Science Foundation. (2014). *Science and engineering indicators.* Washington, D.C.

Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision, 1*(1), 64-78.

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology*, *108*, 802-822.

Peters, E., Lipkus, I., & Diefenbach, M. A. (2006). The functions of affect in health communications and in the construction of health preferences. *Journal of Communication*, *56(s1)*, S140-S162.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*(5), 407-413.

Petrova, D., Garcia-Retamero, R., & Cokely, E. (2015). Understanding harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making,35*(7), 847-858.

Petrova, D., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making, 27*, 191-199.

Reyna, V.F. (2008). A theory of medical decision making and health: Fuzzy trace theory. *Medical Decision Making,28*(6), 850–865.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making*, *7*(3), 332-359.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943-973.

Rimer, B. K., Briss, P. A., Zeller, P. K., Chan, E. C., & Woolf, S. H. (2004). Informed decision making: What is its role in cancer screening? *Cancer, 101*(S5), 1214-1228.

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education & Behavior*, *2(4),* 328-335.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science, 12*(3), 185-190.

Schwartz, P. H. (2011). Decision aids, prevention, and the ethics of disclosure. *Hastings Center Report*, *41*(2), 30-39.

Schwartz, L. M., Woloshin, S., Fowler, F. J., & Welch, H. (2004). Enthusiasm for cancer screening in the United States. *JAMA, 291*(1), 71-78.

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2005). Can patients interpret health information? An assessment of the medical data interpretation test. *Medical Decision Making, 25*(3), 290-300.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*(11), 966-972.

Schwartz, L. M., Woloshin, S., Sox, H. C., Fischhoff, B., & Welch, H. G. (2000). US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: Cross sectional survey. *BMJ, 320*(7250), 1635-1640.

Sheridan, S. L., Harris, R. P., & Woolf, S. H. (2004). Shared decision making about screening and chemoprevention: A suggested approach from the US preventive services task force. *American Journal of Preventive Medicine, 26*(1), 56-66.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24*(2), 311-322.

Slovic, P., Peters, E., Finucane, M. L., & MacGregor, D. G. (2005). Affect, risk, and decision making. *Health Psychology, 24*(4), S35-S40.

Strech, D. (2014). Participation rate or informed choice? Rethinking the European key performance indicators for mammography screening. *Health Policy*, *115*(1), 100-103.

Sutkowi-Hemstreet, A., Vu, M., Harris, R., Brewer, N. T., Dolor, R. J., & Sheridan, S. L. (2015). Adult patients' perspectives on the benefits and harms of overused screening tests: a qualitative study. *Journal of General Internal Medicine*, Epub ahead of print.

van den Berg, M., Timmermans, D. R., Leo, P., van Vugt, J. M., & van der Wal, G. (2006). Informed decision making in the context of prenatal screening. *Patient Education and Counseling, 63*(1), 110-117.

Waller, J., Douglas, E., Whitaker, K. L., & Wardle, J. (2013). Women's responses to information about overdiagnosis in the UK breast cancer screening programme: A qualitative study. *BMJ Open, 3*(4), e002703.

Waller, J., Osborne, K., & Wardle, J. (2015). Enthusiasm for cancer screening in Great Britain: a general population survey. *British Journal of Cancer*, *112*(3), 562-566.

Wegwarth, O., & Gigerenzer, G. (2011). "There is nothing to worry about": Gynecologists' counseling on mammography. *Patient Education and Counseling, 84*(2), 251-256.

Woloshin, S., Schwartz, L. M., Black, W. C., & Kramer, B. S. (2012). Cancer screening campaigns - getting past uninformative persuasion. *New England Journal of Medicine, 367*(18), 1677-1679.

Zikmund-Fisher, B. J. (2012). The right tool is what they need, not what we have: a taxonomy of appropriate levels of precision in patient risk communication. *Medical Care Research and Review*, *70*(1), 37S-49S.

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2010). Risky feelings: Why a 6% risk of cancer does not always feel like 6%. *Patient Education and Counseling, 81*, S87-S93.

CHAPTER 3:

Communication of benefits and harms to patients

The content of this chapter has been submitted as Petrova, D., Garcia-Retamero, R., Delaney, B., Cokely, E., & Kostopoulou, O. Strengths and gaps in physicians' risk communication about screening: the role of numeracy. The online supplement and appendix are found in this link: https://db.tt/rceKvTEZ.

Chapter 3

Strengths and gaps in physicians' risk communication about screening: the role of numeracy

Many patients have low numeracy, which impedes their understanding of important information about health (e.g., benefits and harms of screening). We investigated whether physicians adapt their risk communication to accommodate the needs of patients with low numeracy, and how physicians' own numeracy influences their understanding and communication of screening statistics. UK family physicians (N=151) read a vignette describing a patient seeking advice on cancer screening. In the vignette, we manipulated the numeracy of the patient (low vs. high vs. unspecified). We measured physicians' risk communication, recommendation to the patient, understanding of screening statistics, and numeracy. Consistent with best practices, family physicians generally preferred to use visual aids ($\chi 2(2)=30.14$, $p<.001$) rather than numbers when communicating information to a patient with low (vs. high) numeracy. However, physicians with high (vs. low) numeracy offered more meaningful and complete risk information: they were more likely to mention mortality rates, OR=8.55 [95% CI 1.77, 41.41], $p=.007$, and harms from overdiagnosis, OR=8.82 [1.34, 60.25], $p=.023$. Physicians with high numeracy were also more likely to understand that increased survival rates do not imply screening effectiveness, OR=6.05 [1.27, 28.72], $p=.026$. Many well-intentioned physicians are prone to misunderstanding risks and communicating suboptimal information to their patients. Although less numerate physicians know how to make risks easier to understand, they may not realize how much they themselves can misunderstand and mislead. High-quality risk communication and shared decision making can depend critically on factors that shape the risk literacy of physicians (e.g., numeracy, visual aids).

1. Introduction

Cancer screening can save lives but can also carry risks such as false positive results and the risk of unnecessary treatment (Harris et al., 2014). Given the evidence for potential harm across a number of screening tests for cancer (e.g., breast, prostate, lung, thyroid, Esserman, Thompson, & Reid, 2013), it is likely that informed rather than persuasion-based decision making will become the standard for many screening decisions (Esserman et al., 2013; Sheridan, Harris, & Woolf, 2004; Woloshin, Schwartz, Black, & Kramer, 2012). However, several obstacles to informed decision making have been documented. For example, many US and European adults believe that cancer screening is almost always beneficial. In turn, they tend to grossly overestimate the benefits and are often unaware of possible harms (Gigerenzer, Mata, & Frank, 2009; Hersch et al., 2013; Hoffman et al., 2010; Hudson, Zarifeh, Young, & Wells, 2012; Schwartz, Woloshin, Sox, Fischhoff, & Welch, 2000; Schwartz, Woloshin, Fowler, & Welch, 2004; Waller, Osborne, & Wardle, 2015). People also often perceive that screening is an obligation to family and society and that foregoing screening is irresponsible (Schwartz et al., 2004; Waller et al., 2015). Such beliefs are on occasion reinforced by campaigns that use misleading statistics, exaggerate the benefits, and omit serious harms (Gigerenzer, 2014; Woloshin & Schwartz, 2012). Highly positive attitudes towards cancer screening can stem from the perceived value of saving a life, and the effectiveness of some screening programs; however, such attitudes can also lead to misinformed decisions about screening programs with disputed or mixed efficacy (Arkes & Gaissmaier, 2012; Petrova, Garcia-Retamero, & Cokely, 2015). For example, the United States Preventative Services Task Force recommends against screening for prostate cancer with PSA tests (Moyer, 2012). However, even when shown clearly presented evidence that this screening causes harm and is on average ineffective, some men fail to understand the evidence and are willing to get screened (Petrova et al., 2015).

Research shows that physicians' knowledge of screening benefits and harms is also often incomplete or evolving, and so physicians may fail to discuss harms with their patients (Elstad et al., 2015; Elstad et al., 2015; Han et al., 2013; Hoffman et al., 2010; Wegwarth & Gigerenzer, 2011; Wegwarth & Gigerenzer, 2013). A national survey of US physicians showed that most primary care physicians mistakenly interpreted increased detection and improved survival as evidence that screening saves lives (Wegwarth, Schwartz, Woloshin, Gaissmaier, & Gigerenzer, 2012). This shows a common lack of awareness that survival statistics are distorted by both lead time bias and overdiagnosis bias, and thus do not provide compelling evidence for screening benefits (Wegwarth et al., 2012; Welch, Schwartz, & Woloshin, 2000).

These circumstances highlight the need for careful communication of the risk of harms and benefits from screening. However, despite a growing literature on risk comprehension and risk literacy, there is little direct evidence examining the factors that foster better understanding and communication in this specific context. One factor that could strongly influence communication and understanding in clinical settings is numeracy of physicians and patients. Statistical numeracy in particular refers to the ability to understand and evaluate numerical expressions of probabilities and risk, and is robustly related to superior decision making and risk literacy across many medical and other decision contexts (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Cokely & Kelley, 2009; Garcia-Retamero & Galesic, 2013; Garcia-Retamero, Andrade, Sharit, & Ruiz, 2015; Reyna, Nelson, Han, & Dieckmann, 2009). For example, patients with low numeracy tend to overestimate their risk of cancer and the benefit of medical treatments, and are less capable of using numerical information to inform their perceptions of risks and benefits (Davids, Schapira, McAuliffe, & Nattinger, 2004; Lipkus, Peters, Kimmick, Liotcheva, & Marcom, 2010; Schwartz, Woloshin, Black, & Welch, 1997; Woloshin, Schwartz, Black, & Welch, 1999). While research on patient numeracy is abundant, research in health professionals is much less extensive. Higher physician numeracy is associated with more optimal recommendations regarding Medicare D plans (Hanoch, Miron-Shatz, Cole, Himmelstein, & Federman, 2010), better inferences about screening test results or risks of side effects (Garcia-Retamero, Cokely, Wicki, & Joeris, 2016; Garcia-Retamero & Hoffrage, 2013), and preference for communicating numerical information (Anderson, Obrecht, Chapman, Driscoll, & Schulkin, 2011). However, the influence of physician numeracy on risk communication approaches with patients is unknown. It has been suggested that high physician numeracy should provide risk communication and shared decision making benefits via better understanding of screening statistics and evidence-based recommendations (Garcia-Retamero et al., 2014).

It has also been suggested that physicians adapt their risk communication to the patient's numeracy (Elwyn, Edwards, Kinnersley, & Grol, 2000; Gaissmaier, Anderson, & Schulkin, 2014; Galesic & Garcia-Retamero, 2011; Garcia-Retamero et al., 2014; Hanoch, Miron-Shatz, Rolison, Omer, & Ozanne, 2014; Schwartz, 2011). For example, patients with low numeracy prefer verbal as opposed to numerical information which they have trouble understanding (Fagerlin et al., 2007). Visual aids like icon arrays have gained popularity because they often increase comprehension among patients with low numeracy (Garcia-Retamero & Cokely, 2013). Several instruments exist that can quickly assess patients' health literacy and numeracy skills in primary care settings (Cokely et al., 2012; Weiss et al., 2005). However, it is not known how information about the patient's numeracy influences physicians' risk communication. It is possible that physicians adapt their communication style to the patient's numeracy in accord with best practices, and thus should be *more* likely to use words and visual aids and *less* likely to use numbers, when talking to patients with low numeracy.

Finally, the presence of an official clinical guideline that recommends cancer screening may mean that physicians are more likely to recommend screening. Research suggests that some physicians order screening for their patients even when they do not believe the screening is life-saving. Theoretically, they do so because of strong patient demand, fear of lawsuits, or the belief that it represents the standard of practice (Austin, Valente, Hasse, & Kues, 1997; Hicks, Hamm, & Bemben, 1995; Voss & Schectman, 2001). Accordingly, we investigated three possible determinants of physicians' risk communication and recommendations, namely: (1) their own numeracy, (2) the patient's numeracy, and (3) the presence of an official clinical guideline recommending screening.

2. Method

We conducted an anonymous online survey of family physicians in the UK, a country where screening is offered within national programs by sending invitations to eligible participants with information on benefits and risks (Waller et al., 2015). While family physicians are not directly involved in the delivery of most screening programs, they are usually the first point of contact for patients and may receive inquiries about screening. Potential participants were invited by e-mail to complete a 15-minute survey about "communication of information regarding cancer screening to patients", in return for a £10 Amazon voucher. E-mail addresses were obtained from a database of physicians who had participated in previous studies of one of the authors (OK). In October 2015, we emailed a total of 516 board certified and currently practising family physicians. Follow-up emails were sent to non-respondents until the required sample size was reached. We used G*Power to calculate the required sample size for regression analyses to detect a small effect size (f^2=0.09) with an alpha of 0.05, a power of 80%, and 5 predictors. The minimum sample size required was 149.

2.1. Design, materials and procedure

Participants first responded to standard demographic questions (age, gender, year of medical degree, years of experience in family practice, type of family practice). Subsequently, participants read a brief vignette about 'Sam', a fictitious 61-year old patient, who came for advice regarding a screening test for cancer X. No specific cancer was mentioned to avoid the influence of participants' knowledge about existing cancer screening programs. The depicted benefits and risks were designed to be realistic and plausible given the current state of the science of cancer screening. The vignette and associated questions were developed with the help of an experienced family physician (BD), and were revised after pilot testing with two other family physicians. We created 12 versions of the vignette by manipulating the presence of a clinical guideline (present vs. absent), the effectiveness of the screening test (effective vs. not effective), and patient numeracy (low vs. high vs. not specified). Participants were randomized to one of the 12 vignette versions. Depending on the version, participants read information about the clinical guideline, screening effectiveness, and patient numeracy, as described in Figure 1. After reading the vignette, participants answered a series of questions measuring the following:

1a. Risk communication quality. Participants were shown a list of topics and asked to indicate which ones they would discuss with the patient: (1) detection of cancer X, (2) mortality from cancer X, (3) false positive screening tests, and (4) number of people overdiagnosed and treated unnecessarily, with and without screening. Based on the topics that they selected, we created a quality index based on published guidelines for completeness and interpretability of risk communication (see Appendix for details) (W.

Gaissmaier et al., 2014). The index ranged from 0 (low quality) to 4 (high quality).

1b. Risk communication format. For each topic selected above, participants were presented with a list of options: conveying the information with words (e.g., small, larger, etc.), visual aids (e.g., draw a simple graph, use an icon array), and numbers (e.g., percentages, number of people). Participants could choose more than one communication format. We counted how many times they chose each format.

2. Recommendation. Participants indicated their screening recommendation to the patient by choosing either "in favour", or "against", or "neutral" (neither for nor against).

3. Understanding of cancer screening statistics. We administered a short questionnaire by Wegwarth et al.(O. Wegwarth et al., 2012). For each of the three statements presented in Table 2, physicians indicated whether it proved or not that screening saves lives. We computed the sum of correct answers (range 0 to 3).

Physician numeracy. This was measured with the Berlin Numeracy Test-Schwartz (BNT-S) following Cokely et al. (E. T. Cokely et al., 2012) (see RiskLiteracy.org). The score ranged from 1 to 7.

The full survey is available at http://tinyurl.com/UKGPSurvey. [6]

2.2. Analysis

To measure the effect of the manipulated factors and physician numeracy on the outcome measures, we conducted multiple regressions using the GENLIN procedure in SPSS 21. Specifically, we conducted Poisson regressions for the risk communication format measures, which were simple counts of the

---

[6] We additionally measured physicians' usual and preferred roles in decision making about screening. Higher physician numeracy was related to more patient involvement in cancer screening decisions, consistent with previous research. (R. Garcia-Retamero, Wicki, Cokely, & Hanson, 2014) These complementary results are reported in an online appendix: http://tinyurl.com/UKGPstudy.

number of times participants selected each format. In these regressions we also controlled for the total number of topics chosen. We conducted ordinal regressions for all other outcome measures, which were ordinal. We computed odds ratios for physician numeracy representing one unit change on the 1-7 scale of measurement: OR. We also computed odds ratios for physician numeracy that represented a six-unit change on the1-7 scale, in order to provide a more intuitive measure of effect size, comparing the odds of physicians with the highest vs. low numeracy: $OR_{min-max}$.

3. Results

Of 516 invited family physicians, 174 (34%) started the survey and 151 (29%) completed it. Half of the participants were female and the majority (52%) worked in urban practices. More characteristics of the sample and descriptive statistics for the measures are found in Tables 1 to 3.

1a. Risk communication quality. Physician numeracy was positively related to the quality index of risk communication (range 0-4): OR=1.39 [1.10, 1.76], $OR_{min-max}$=7.21 [1.77, 29.72], *p*=.007. When we performed logistic regressions on each of the topics that physicians could choose from to discuss with the patient, we found that increasing physician numeracy was associated with increasing odds of discussing mortality rates (OR=1.43 [1.10, 1.86], $OR_{min-max}$=8.55 [1.77, 41.41], *p*=.007), and harms from overdiagnosis (OR=1.44 [1.05, 1.98], $OR_{min-max}$=8.82 [1.34, 60.25], *p*=.023). Patient numeracy had no effect on the quality of risk communication (*p*=.985).

1b. Risk communication format. If the patient was described as having low (vs. high or unspecified) numeracy, physicians more often chose to communicate risk using visual aids (Wald $\chi^2(2)$=30.14, *p*<.001), and less often using numbers (Wald $\chi^2(2)$=15.71, *p*<.001, see Table 1S in the online supplement). Patient numeracy had no significant effect on how often physicians chose words, *p*=.529. Physician numeracy had no effect on format (*p*=.198 for visual aids, *p*=.665 for numbers, and *p*=.990 for words), and did not interact with patient numeracy (*p*=.557 for visual aids, *p*=.134 for numbers, and *p*=.317 for words).

2. Recommendation. Thirty-five (23%) physicians recommended screening to the patient, 57 (38%) recommended not to get screened, and 59 (39%) were neutral. Physicians tended to recommend screening more often when there was a clinical guideline (27%) than when there was not (19%): OR=1.68 [95% CI 0.92-3.11], *p*=.096. Twenty percent of physicians recommended screening when it was not effective (vs. 27% who recommended it when it was effective) but the difference was borderline (*p*=.062). Patient numeracy (*p*=.076), physician numeracy (*p*=.845), and its interaction with screening effectiveness (*p*=.204) had no effect on recommendations.

3. Understanding of screening statistics. Physician numeracy was positively related to the understanding of screening statistics (range 0-3): OR=1.61 [1.26, 2.06], $OR_{min-max}$=17.41 [4.01, 76.42], *p*<.001. With increasing numeracy, physicians were more likely to know that neither increased detection (OR=1.71 [1.25, 2.34], $OR_{min-max}$=25.00 [3.81, 164.17], *p*=.001), nor survival rates (OR=1.35 [1.04, 1.75], $OR_{min-max}$=6.05 [1.27, 28.72], *p*=.026), demonstrated screening effectiveness. Numeracy was not associated with knowing that decreased mortality found in a randomized trial demonstrated screening effectiveness (*p*=.151).

Table 1. Characteristics of the sample and descriptive statistics for the dependent measures. N=151. SD=standard deviation.

	Minimum	Maximum	Mean	SD
Physician age	30	65	42.3	8.2
Physician experience: number of years in family practice	2	35	12.2	8.3
Physician numeracy (the Berlin Numeracy Test-Schwartz [24])	1	7	5.7	1.3
Risk communication quality index	0	4	2.8	1.3
Total number of risk communication topics selected	1	8	4.8	1.9
Risk communication format: choosing words	0	8	1.3	1.9
Risk communication format: choosing visual aids	0	8	1.2	2.0
Risk communication format: choosing numbers	0	8	2.3	2.5
Understanding of screening statistics score	0	3	1.9	.8

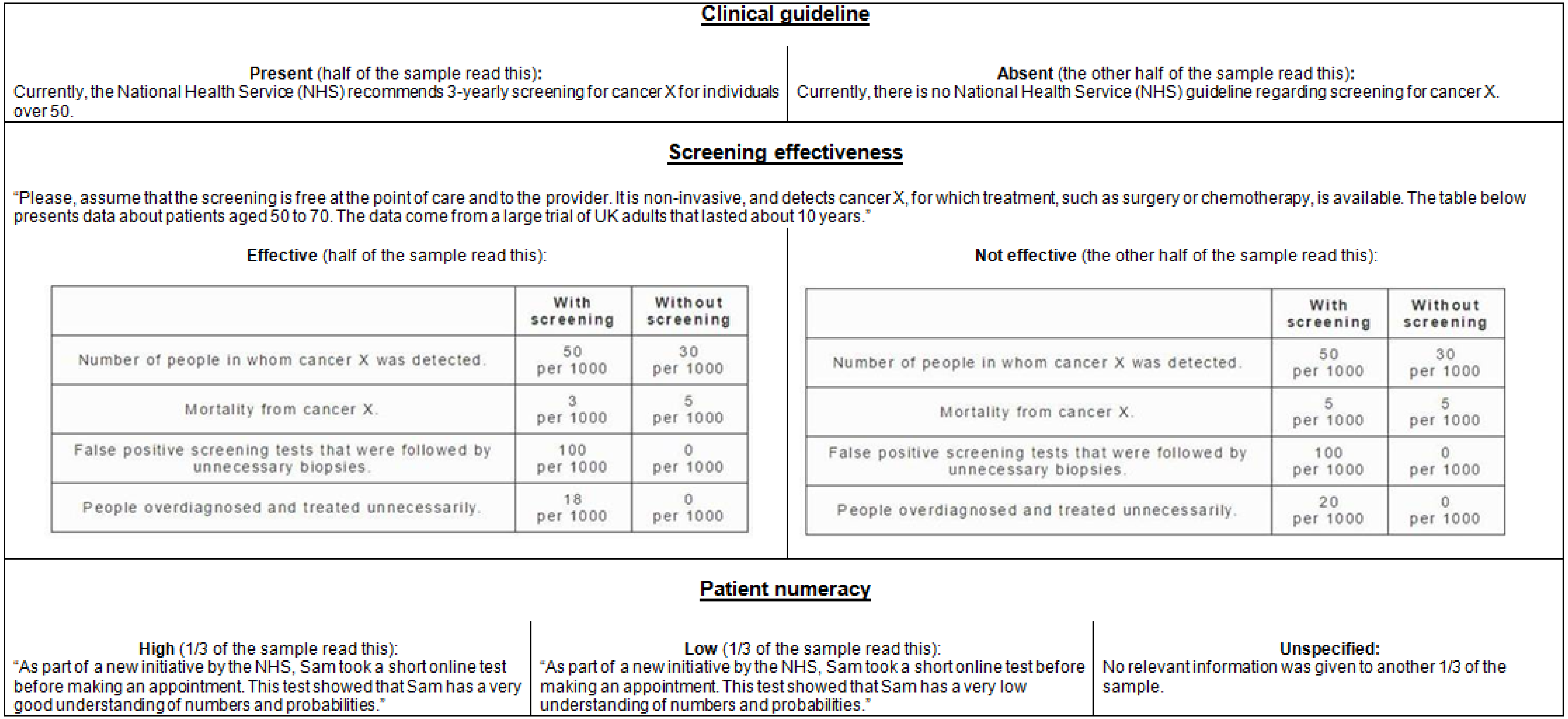
Table 2. Number of physicians (% of total sample - N=151) that chose to discuss each of eight topics with the patient.

	With screening	Without screening
Detection of cancer X	113 (75)	96 (64)
Mortality from cancer X	103 (68)	104 (69)
False positive screening tests	119 (79)	32 (21)
People overdiagnosed and treated unnecessarily	125 (83)	37 (25)

Table 3. Items used to assess physician understanding of cancer screening statistics from Wegwarth et al.(2012): numbers of physicians (% of the total sample - N=151). The correct answer to each question is marked with an asterisk.

	Proves	Does not prove	I don’t know
More cancers are detected in screened populations than in unscreened populations.	24 (16)	121 (80)*	6 (4)
Cancers detected at screening have better 5-year survival rates than cancers detected because of symptoms.	77 (51)	64 (42)*	10 (7)
In a randomised trial, mortality rates of screened persons are lower than those of unscreened persons.	100 (66)*	36 (24)	15 (10)

Figure 1. Information provided to participants based on vignette version. Note: The statistics shown to communicate screening effectiveness were fictitious but based on outcomes of some common cancer screenings (e.g., screening for prostate cancer with PSA tests and screening for breast cancer with mammography, Gøtzsche & Jørgensen, 2013; Ilic, Neuberger, Djulbegovic, & Dahm, 2013).



Clinical guideline

Present (half of the sample read this)**:**
Currently, the National Health Service (NHS) recommends 3-yearly screening for cancer X for individuals over 50.

Absent (the other half of the sample read this):
Currently, there is no National Health Service (NHS) guideline regarding screening for cancer X.

Screening effectiveness

"Please, assume that the screening is free at the point of care and to the provider. It is non-invasive, and detects cancer X, for which treatment, such as surgery or chemotherapy, is available. The table below presents data about patients aged 50 to 70. The data come from a large trial of UK adults that lasted about 10 years."

Effective (half of the sample read this):

	With screening	Without screening
Number of people in whom cancer X was detected.	50 per 1000	30 per 1000
Mortality from cancer X.	3 per 1000	5 per 1000
False positive screening tests that were followed by unnecessary biopsies.	100 per 1000	0 per 1000
People overdiagnosed and treated unnecessarily.	18 per 1000	0 per 1000

Not effective (the other half of the sample read this):

	With screening	Without screening
Number of people in whom cancer X was detected.	50 per 1000	30 per 1000
Mortality from cancer X.	5 per 1000	5 per 1000
False positive screening tests that were followed by unnecessary biopsies.	100 per 1000	0 per 1000
People overdiagnosed and treated unnecessarily.	20 per 1000	0 per 1000

Patient numeracy

High (1/3 of the sample read this):
"As part of a new initiative by the NHS, Sam took a short online test before making an appointment. This test showed that Sam has a very good understanding of numbers and probabilities."

Low (1/3 of the sample read this):
"As part of a new initiative by the NHS, Sam took a short online test before making an appointment. This test showed that Sam has a very low understanding of numbers and probabilities."

Unspecified:
No relevant information was given to another 1/3 of the sample.

4. Discussion

Compared to physicians with low numeracy, physicians with the highest numeracy in our sample had 7 times the odds of offering complete and meaningful risk communication about screening to patients. Higher numeracy was related to greater likelihood of communicating (1) mortality rates to the patient, information that is highly relevant for evaluating screening benefits (Wegwarth et al., 2012; Welch et al., 2000), and (2) risks of unnecessary treatment, a harm that is still largely unknown to the public (Schwartz et al., 2000; 2004). Physicians with higher numeracy were also less likely to think that increased survival rates from screening demonstrate that screening saves lives (Wegwarth et al., 2012). Overall, these results clearly show the benefits of physician numeracy for fostering risk literacy and evidence-based decisions in clinical practice. Results also suggest that physicians with low numeracy may generally be ill-equipped to achieve informed decision making about cancer screening. A previous study linked physicians' lower data interpretation abilities (e.g., distinguishing between relative and absolute risk) to more enthusiasm about cancer screening in general (Caverly, Prochazka, Binswanger, Kutner, & Matlock, 2014). While we did not find a relationship between physician numeracy and screening recommendation, our findings do suggest that physicians with low numeracy are much more likely to provide patients with insufficient and lower quality information about screening.

It is encouraging that physicians successfully adapted their risk communication to suit the patient's level of numeracy. There is evidence that visual aids enhance the understanding of risk by patients with low numeracy (Garcia-Retamero & Cokely, 2013). Our participants' choice of visual aids as the most preferred mode of communicating risk to low numeracy patients suggests that physicians should have the necessary tools to assess patient numeracy where appropriate, and are both trained and willing to use visual aids for communicating screening benefits and harms. Even if physicians don’t know that visual aids make risks easier to comprehend, these powerful and simple tools would be a ready risk communication means at their disposal. However, the presence of an official guideline was only related to a small increase in screening recommendations, regardless of the evidence for benefit. This is in line with previous findings showing that physicians sometimes order screening for their patients because they believe this is the standard of practice, even if they are not convinced that screening improves patient outcomes (Voss & Schectman, 2001).

Because the current study only used a single, hypothetical scenario we may not have sufficiently simulated the difficulties and nuances of the doctor-patient communication, including the pressure from patient demands. Although the current results are suggestive, the extent to which these risk communication approaches appear in every practice remains an open question. Going forward, the influence of physician numeracy should also be further explored in information exchanges between physicians and patients, and in relation to actual cancer screening decisions.

5. Conclusion

Numeracy is an essential component of risk literacy that tends to cultivate informed, evidence-based decisions (Cokely et al., 2012; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007). Unfortunately, not all patients and physicians have developed the skills that tend to be necessary for independently evaluating and understanding the complex statistics that are now part of many medical decisions. On one hand, this research suggests that screening patients for numeracy may help many physicians tailor risk communication to patient needs and abilities, partially mitigating this problem (Trevena et al., 2013; Zikmund-Fisher, 2013). On the other hand, this research shows how insufficient physician numeracy can impede informed decision making about screening, because as is common among many diverse professionals, many physician's simply do not know what they do not understand (Ghazal, Cokely, & Garcia-Retamero, 2014). In our study about 20% of physicians recommended that a hypothetical patient participate in screening that was not effective and that was likely to cause serious harm. To avoid such misguided recommendations, it is necessary to improve numeracy, risk literacy, and statistical skills training in medical curricula and continuing education (Gigerenzer et al., 2007; Schulkin & Anderson, 2014; Rao, 2008). Beyond this, our results also suggest there is a pressing need for well-designed decision aids to help less numerate physicians and patients understand and discuss life-altering risks and benefits. Examples are simple facts boxes that display the most important information in a tabular format, accompanied by visual aids in the form of icon arrays (Arkes & Gaissmaier, 2012; Schwartz, Woloshin, & Welch, 2007; Zikmund-Fisher, Fagerlin, & Ubel, 2008; see Harding Center for Risk Literacy (www.harding-center.mpg.de) for excellent examples of risk communication materials for several cancer screenings). Such decisions aids can be easily implemented in clinical practice and can effectively increase comprehension among both patients and physicians (Garcia-Retamero et al., 2016; Petrova et al., 2015). Comprehension in turn can help patients and physicians discuss and evaluate risks and benefits, in light of the patient's values and informed by the physician's expertise, laying the grounds for shared decision making.

6. References

Anderson, B. L., Obrecht, N. A., Chapman, G. B., Driscoll, D. A., & Schulkin, J. (2011). Physicians' communication of Down syndrome screening test results: The influence of physician numeracy. *Genetics in Medicine, 13*(8), 744-749.

Arkes, H. R., & Gaissmaier, W. (2012). Psychological research and the prostate-cancer screening controversy. *Psychological Science, 23*(6), 547-553.

Austin, O. J., Valente, S., Hasse, L. A., & Kues, J. R. (1997). Determinants of prostate-specific antigen test use in prostate cancer screening by primary care physicians. *Archives of Family Medicine, 6*(5), 453-458.

Caverly, T. J., Prochazka, A. V., Binswanger, I. A., Kutner, J. S., & Matlock, D. D. (2014). Confusing relative risk with absolute risk is associated with more enthusiastic beliefs about the value of cancer screening. *Medical Decision Making, 34*(5), 686-692.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*, 25-47.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4*(1), 20-33.

Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine, 19*(4), 310-315.

Elstad, E. A., Sutkowi-Hemstreet, A., Sheridan, S. L., Vu, M., Harris, R., Reyna, V. F., . . . Brewer, N. T. (2015). Clinicians' perceptions of the benefits and harms of prostate and colorectal cancer screening. *Medical Decision Making, 35*(4), 467-476.

Elwyn, G., Edwards, A., Kinnersley, P., & Grol, R. (2000). Shared decision making and the concept of equipoise: The competences of involving patients in healthcare choices. *The British Journal of General Practice, 50*(460), 892-899.

Esserman, L. J., Thompson, I. M., Jr. Reid, B. (2013). Overdiagnosis and overtreatment in cancer: An opportunity for improvement. *JAMA, 310*(8), 797-798.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test:

Development of the Subjective Numeracy Scale. *Medical Decision Making, 27*(5), 672-680.

Gaissmaier, W., Anderson, B. L., & Schulkin, J. (2014). How do physicians provide statistical information about antidepressants to hypothetical patients? *Medical Decision Making, 34*(2), 206-215.

Galesic, M., & Garcia-Retamero, R. (2011). Do low-numeracy people avoid shared decision making? *Health Psychology, 30*(3), 336-341.

Garcia-Retamero, R., Cokely, E. T., Wicki, B., & Joeris, A. (2016). Improving risk literacy in surgeons. *Patient Education and Counseling,* Advance online publication.

Garcia-Retamero, R., & Galesic, M. (2013). *Transparent communication of health risks: Overcoming cultural differences*. New York: Springer.

Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychology, 33*(8), 920-928.

Garcia-Retamero, R., Andrade, A., Sharit, J., & Ruiz, J. G. (2015). Is patients' numeracy related to physical and mental health? *Medical Decision Making, 35*(4), 501-511.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22*(5), 392-399.

Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine, 83*(0), 27-33.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women. *British Medical Journal, 348,* g2636.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest, 8*(2), 53-96.

Gigerenzer, G., Mata, J., & Frank, R. (2009). Public knowledge of benefits of breast and prostate cancer screening in Europe. *Journal of the National Cancer Institute, 101*(17), 1216-1220.

Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *Cochrane Database of Systematic Revews, 6,* CD001877.

Han, P. K. J., Kobrin, S., Breen, N., Joseph, D. A., Li, J., Frosch, D. L., & Klabunde, C. N. (2013). National evidence on the use of shared decision making in prostate-specific antigen screening. *The Annals of Family Medicine, 11*(4), 306-314.

Hanoch, Y., Miron-Shatz, T., Cole, H., Himmelstein, M., & Federman, A. D. (2010). Choice, numeracy, and physicians-in-training performance: The case of Medicare Part D. *Health Psychology, 29*(4), 454-459.

Hanoch, Y., Miron-Shatz, T., Rolison, J. J., Omer, Z., & Ozanne, E. (2014). Shared decision making in patients at risk of cancer: The role of domain and numeracy. *Health Expectations, 18*(6), 2799–2810.

Harris, R. P., Sheridan, S. L., Lewis, C. L., Barclay, C., Vu, M. B., Kistler, C. E., . . . Brewer, N. T. (2014). The harms of screening: A proposed taxonomy and application to lung cancer screening. *JAMA Internal Medicine, 174*(2), 281-286.

Hersch, J., Jansen, J., Barratt, A., Irwig, L., Houssami, N., Howard, K., . . . McCaffery, K. (2013). Women's views on overdiagnosis in breast cancer screening: A qualitative study. *BMJ, 346*, f158.

Hicks, R. J., Hamm, R. M., & Bemben, D. A. (1995). Prostate cancer screening. What family physicians believe is best. *Archives of Family Medicine, 4*(4), 317-322.

Hoffman, R. M., Lewis, C. L., Pignone, M. P., Couper, M. P., Barry, M. J., Elmore, J. G., . . . Zikmund-Fisher, B. J. (2010). Decision-making processes for breast, colorectal, and prostate cancer screening: The DECISIONS survey. *Medical Decision Making, 30*(5 Suppl), 53S-64S.

Hudson, B., Zarifeh, A., Young, L., & Wells, J. E. (2012). Patients' expectations of screening and preventive treatments. *Annals of Family Medicine, 10*(6), 495-502.

Ilic, D., Neuberger, M. M., Djulbegovic, M., & Dahm, P. (2013). Screening for prostate cancer. *Cochrane Database of Systematic Reviews, 1,* CD004720.

Schulkin, J., & Anderson, B. L. (2014). Physicians’ understanding and use of numeric information. In B. L. Anderson & J. Schulkin (Eds.), *Numerical Reasoning in Judgments and Decision Making about Health* (pp. 59-79). Cambridge University Press.

Lipkus, I. M., Peters, E., Kimmick, G., Liotcheva, V., & Marcom, P. (2010). Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Medical Decision Making, 30*(4), 464-473.

Schwartz, L. M., Woloshin, S., Sox, H. C., Fischhoff, B., & Welch, H. G. (2000). US women's attitudes to false positive mammography results and detection of ductal carcinoma in situ: Cross sectional survey. *BMJ, 320*(7250), 1635-1640.

Moyer, V. A. (2012). Screening for prostate cancer: US preventive services task force recommendation statement. *Annals of Internal Medicine, 157*(2), 120-134.

Petrova, D., Garcia-Retamero, R., & Cokely, E. T. (2015). Understanding the harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making, 35*(7), 847-858.

Rao, G. (2008). Physician numeracy: Essential skills for practicing evidence-based medicine. *Family Medicine, 40*(5), 354-358.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943-973.

Schwartz, L. M., Woloshin, S., Fowler Jr, F. J., & Welch, H. G. (2004). Enthusiasm for cancer screening in the United States. *JAMA*, *291*(1), 71-78.

Schwartz, P. H. (2011). Decision aids, prevention, and the ethics of disclosure. *Hastings Center Report, 41*(2), 30-39.

Schwartz, L. M., Woloshin, S., & Welch, H. G. (2007). The drug facts box: Providing consumers with simple tabular data on drug benefit and harm. *Medical Decision Making, 27*(5), 655-662.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*(11), 966-972.

Sheridan, S. L., Harris, R. P., & Woolf, S. H. (2004). Shared decision making about screening and chemoprevention: A suggested approach from the US Preventive Services Task Force. *American Journal of Preventive Medicine, 26*(1), 56-66.

Trevena, L. J., Zikmund-Fisher, B. J., Edwards, A., Gaissmaier, W., Galesic, M., Han, P. K., . . . Lipkus, I. (2013). Presenting quantitative information about decision outcomes: A risk communication primer for patient decision aid developers. *BMC Medical Informatics and Decision Making, 13*(Suppl 2), S7.

Voss, J. D., & Schectman, J. M. (2001). Prostate cancer screening practices and beliefs. *Journal of General Internal Medicine, 16*(12), 831-837.

Waller, J., Osborne, K., & Wardle, J. (2015). Enthusiasm for cancer screening in Great Britain: A general population survey. *British Journal of Cancer, 112*(3), 562-566.

Wegwarth, O., & Gigerenzer, G. (2011). "There is nothing to worry about": Gynecologists' counseling on mammography. *Patient Education and Counseling, 84*(2), 251-256.

Wegwarth, O., & Gigerenzer, G. (2013). Overdiagnosis and overtreatment: Evaluation of what physicians tell their patients about screening harms. *JAMA Internal Medicine, 173*(22), 2086-2087.

Wegwarth, O., Schwartz, L. M., Woloshin, S., Gaissmaier, W., & Gigerenzer, G. (2012). Do physicians understand cancer screening statistics? A national survey of primary care physicians in the united states. *Annals of Internal Medicine, 156*(5), 340-349.

Weiss, B. D., Mays, M. Z., Martz, W., Castro, K. M., DeWalt, D. A., Pignone, M. P., . . . Hale, F. A. (2005). Quick assessment of literacy in primary care: The newest vital sign. *Annals of Family Medicine, 3*(6), 514-522.

Welch, H. G., Schwartz, L. M., & Woloshin, S. (2000). Are increasing 5-year survival rates evidence of success against cancer? *JAMA, 283*(22), 2975-2978.

Woloshin, S., & Schwartz, L. M. (2012). How a charity oversells mammography. *BMJ, 345*, e5132.

Woloshin, S., Schwartz, L. M., Black, W. C., & Welch, H. G. (1999). Women's perceptions of breast cancer risk: How you ask matters. *Medical Decision Making, 19*(3), 221-229.

Woloshin, S., Schwartz, L. M., Black, W. C., & Kramer, B. S. (2012). Cancer screening campaigns — getting past uninformative persuasion. *New England Journal of Medicine, 367*(18), 1677-1679.

Zikmund-Fisher, B. J. (2013). The right tool is what they need, not what we have: A taxonomy of appropriate levels of precision in patient risk communication. *Medical Care Research and Review, 70*(1 Suppl), 37S-49S.

Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2008). Improving understanding of adjuvant therapy options by using simpler risk graphics. *Cancer, 113*(12), 3382-3390.

CHAPTER 4:

Promoting informed decision making during a crisis

The content of this chapter has been submitted as Cokely, E. T., Petrova, D., Feltz, A., Ripberger, J. T., Gigerenzer, G., & Garcia-Retamero, R. Promoting informed decision making during a crisis. The appendix and supplemental material are found in this link: https://db.tt/wekaV9n9

Chapter 4

Promoting informed decision making during a crisis

Three weeks after the first confirmed case of Ebola infection in the United States, at the peak of the worst Ebola pandemic in history we conducted a risk communication experiment with diverse U.S. residents testing a model of informed decision making under dread risk. Although messages framed in terms of “Ebola” and “risk” increased fear and confusion, numeracy and honest risk communications featuring visual aids increased deliberation, confidence, and risk comprehension, reducing anxiety and distress. In turn, people who better understood risks rejected policy proposals likely to cost lives (e.g., redirecting funds from cardiovascular care to Ebola treatment), and were less willing to engage in avoidance behaviors (e.g., canceling medical visits, keeping children home from school). Results indicate that even when emotions run hot and message framing biases judgment during an unprecedented national crisis, factors that promote risk literacy empower adaptive and informed decision making.

1. Introduction

"The ability to sense and avoid harmful environmental conditions is necessary for the survival of all living organisms" Paul Slovic (p. 236, 1987)

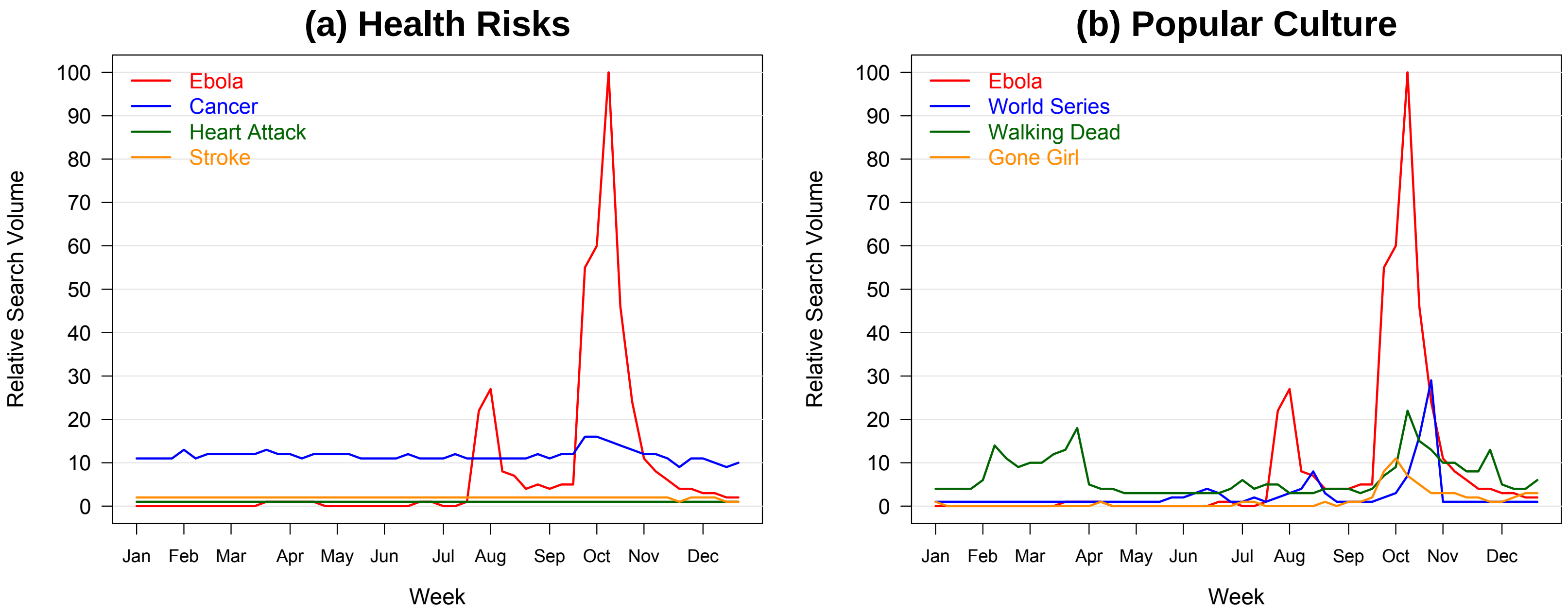
The worst Ebola pandemic in history began in December of 2013, causing more than 11,000 fatalities before the end of 2015. By nearly any standard the Ebola virus is a terrifying and gruesome killer that leads to death in about 50% of those infected. Infections typically begin with flu-like symptoms that are soon followed by extensive bleeding from the eyes, ears, skin, and elsewhere. Over the course of the next two or more weeks, the virus progressively attacks and destroys nearly every tissue in the body. In order to treat and contain the virus, many of those infected are completely isolated from their loved ones as blood loss accelerates and their internal organs fail. The profound suffering caused by Ebola also extends to the families and communities of those infected who endure long-lasting emotional, social, and economic hardships.

Psychologically, Ebola represents a type of dread risk—i.e., a high-consequence low probability event provoking strong emotions and decision biases (Fischhoff, 2013; Gigerenzer, 2014; Morgan, Fischhoff, Bostrom, & Atman, 2002; Slovic, 1987). For example, the dread risks associated with the 9/11 terror attacks in the U.S. led many U.S. residents to choose driving instead of flying, which caused an estimated 1,600 additional deaths in the twelve months following the attacks (Gaissmaier & Gigerenzer, 2012; Gigerenzer, 2006). Although the most recent Ebola pandemic was concentrated in West Africa, on September 30th, 2014, the Center for Disease Control (CDC) confirmed the first ever case of Ebola infection in the United States. The index patient who eventually died from his infection was a man who had recently flown from West Africa to Dallas, Texas. In the course of treating the index patient two brave U.S. healthcare workers also became infected with Ebola. Thankfully, however, both healthcare workers eventually recovered. The infection didn't spread more widely in part because Ebola is relatively hard to contract. While Ebola infection carries a high risk of death (i.e., about 50%), Ebola is only spread through direct contact with infected bodily fluids (e.g., isn't spread by food, air, or water). This low risk of contagion was reflected in initial policies like the CDC's decision not to monitor all the passengers traveling with the index patient on the flight from Africa to Dallas. Nevertheless, the CDC did provide the public with extensive information about the medical risks and consequences of Ebola infection, which are well-understood and have been extensively studied for nearly 50 years.

The dread risks of Ebola infection in the U.S. led to months of vivid media speculation and debate that produced massive spikes in public attention to the disease. Estimates of the relative attention shifts and time courses using google search trend analyses are presented in Figure 1 (Ginsberg, Mohebbi, Patel, Brammer, Smolinski, & Brilliant 2009; Ripberger, 2011; Swearingen and Ripberger, 2014). The largest spike in attention started at the end of September 2014 and peaked between October 12th and 19th, 2014. Trend analyses indicate that during October of 2014 U.S. residents queried "Ebola" up to 100 times as often as they queried other deadly diseases, popular sporting events, and entertainment programs.

As U.S. public concern about Ebola grew, a congressional hearing was convened on October 24th to evaluate major aspects of federal agency responses to Ebola in the United States. That same day states including New York and New Jersey implemented policies that ignored the CDC's evidence-based recommendations. Instead, these States' policies mandated both extensive monitoring and quarantine for those who had close contact with infected individuals including health care workers returning from crucial Ebola containment sites in Liberia, Guinea, and Sierra Leone.

Figure 1. Google searches for "Ebola" compared to (a) other health risks and (b) popular culture. Data are from Google Trends (https://www.google.com/trends/). Figure 1(a) compares the volume of U.S. Google searches for "Ebola" to searches for more significant health risks. For most of 2014, searches for "Cancer," "Heart Attack," and "Stroke" were more prominent than searches for "Ebola." In mid-October, searches for "Ebola" exceeded searches for these risks by factors of 6:1, 100:1, and 50:1, respectively. Figure 1(b) repeats this analysis but compares searches for "Ebola" to searches for the most popular sporting event in October 2014 ("World Series"), a leading (and Halloween-oriented) television series premier ("Walking Dead"), and one of the highest-grossing movies that was released in October 2014 ("Gone Girl"). Again, searches for "Ebola" were relatively uncommon for most of 2014 but exploded in mid-October. At their respective peaks, searches for "Ebola" surpassed searches for the "World Series" by approximately 3:1, "Walking Dead" by 4:1, and "Gone Girl" by a ratio of 9:1.



Theoretically, such quarantines might reduce public risk perceptions thereby protecting the public from their own over-reactions to fears about Ebola. However, many scientists including leaders at the National Institutes for Health stated publically that these new policies may put more U.S. lives at risk because the best way to protect the U.S. from Ebola would be to stop it in West Africa. That is, if mandatory quarantines reduced the number of people willing to travel to help contain Ebola, the pandemic could spread beyond control. Regardless of the merits of the arguments, it stands to reason that public policies are likely to be ineffective when they do not adequately anticipate the capabilities, decisions, or responses of the public. Can people be trusted to make well-informed decisions in the face of the worst Ebola pandemic in history?

Informed Decision Making & Risk Literacy. Widely held standards for informed decision making emphasize the importance of opportunities for deliberation about risk in the light of one's values (Drane, 1984). At the heart of these theories is the notion that an informed decision maker should have a representative but not necessarily extensive understanding of the risks and consequences of various courses of action (Feltz, 2015; Fischhoff, Brewer, & Downs, 2011). Theoretically, one's practical ability to use math in everyday problem solving contexts (i.e., numeracy) should promote informed decision making in many ways. Statistical numeracy in particular is one of the strongest predictors of general decision making skills including one's ability to understand and evaluate information about risk—i.e., risk literacy (www.RiskLiteracy.org) (Bruine de Bruin, Parker, & Fischhoff, 2007; Cokely, Ghazal, Galesic, Garcia-Retamero, & Schulz, 2013; Fagerlin et al., 2007; Ghazal, Cokely, & Garcia-Retamero, 2014; Gigerenzer, 2012; Levin, Bossard, Gaeth, & Yan, 2014; Lipkus, Samsa, & Rimer, 2001; Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008; Parker & Fischhoff, 2005; Peters, Hibbard, Slovic, & Dieckmann, 2007; Peters, Baker, Dieckmann, Leon, & Collins, 2010; Peters, 2012; Reyna, 2004; Reyna, Nelson, Han, & Dieckmann, 2009; Schapira et al., 2012; Schwartz, Woloshin, Black, & Welch, 1997).

Although numeracy is a strong predictor or differences in risk literacy, one's ability to understand and evaluate risk is not simply an individual difference. Social, organizational, and environmental factors can also enhance or reduce risk literacy. For example, transparent risk communications with user-friendly formats have been shown to improve risk literacy among diverse and vulnerable individuals, including people who have very low levels of numeracy (Garcia-Retamero & Galesic, 2013; Garcia-Retamero Cokely, 2013, 2014; Gigerenzer & Hoffrage, 1995; Okan, Garcia-Retamero, Cokely, Maldonado, 2012; Spiegelhalter, Pearson, & Short, 2011; Trevena et al., 2013). In several cases, simple visual aids have been found to improve decision outcomes as much as leading persuasion techniques or extensive educational interventions (Garcia-Retamero Galesic, 2013; Garcia-Retamero & Cokely, 2013).

While there are many ways to help decision makers develop a more thorough understanding of relevant risks, even well-informed decision makers still may not make decisions that accord with their understanding. It is well established that people are not neo-classically rational decision makers who optimize choices by weighting and integrating all available information in abstracted cost/benefit type analyses (Gigerenzer & Gaissmaier, 2011). In contrast, boundedly rational people who have limited time and cognitive capacities tend to rely on simple heuristic strategies that ignore information to facilitate decision making. Even highly numerate individuals rely heavily on simple metacognitive heuristics and potentially bias-inducing affective intuitions during decision making under risk (Cokely & Kelley, 2009; Peters et al., 2006).

Considering the many potential obstacles to informed decision making, and given the unique opportunity to field an experiment at the height of an unprecedented emerging national crisis, we conducted a risk communication experiment with diverse U.S. residents three weeks after the first case of Ebola was confirmed in the United States. Broadly, our aim was to test a theoretical decision process model of factors that promote informed decision making under dread risk, with a specific emphasis on understanding processes that support socially and economically significant decision outcomes. For example, we assessed factors that helped inoculate people against unnecessary and unhealthy avoidance-oriented behavioral intentions such as keeping children home from school, skipping work, and canceling needed medical appointments. Similarly, we mimicked key features of ongoing community debates regarding the merits of new proposed Ebola policies, modeling the influences of skills, decision aids, and various biases on public policy proposals endorsements (e.g., reallocation of funds from programs for cardiovascular disease, HIV/AIDs, Malaria, and Tuberculosis care to U.S. Ebola prevention and treatment).

2. Method

On October 19th, 2014, at the height of public attention to the risks of Ebola we conducted an online risk communication study with 505 adults from around the U.S., recruited for participation via a paid web-panel service. Following informed consent and general instructions, we assessed participants' numeracy levels using the adaptive version of the Berlin Numeracy Test for the general population (i.e., adaptive BNT-S) (Cokely et al., 2012). The experimental protocol followed a modified mixed-factorial design (Figure 2) in which we systematically manipulated three independent variables: (a) the use of the terms "Ebola hemorrhagic fever" vs. "Viral hemorrhagic fever"; (b) the use of the terms "risks" vs. "likelihoods"; and (c) the use of numerical risk information with or without visual aids (Fig. 3). The effects of numeracy and the three experimentally manipulated variables were measured on four dimensions of dread risk-related outcomes: (i)

Comprehension, which included questions about the risk of infection and death from Ebola; (ii) Negative affect, which included questions about being anxious, discouraged, disturbed and worried about Ebola; (iii) Behavioral intentions, which included questions about potential reactions such as whether participants would go to work, use public transportation, and send their children to school; and (iv) Policy endorsements, which included the evaluation of proposals to redirect funds to Ebola prevention and treatment, and away from much more beneficial prevention programs for cardiovascular disease, HIV/AIDS, malaria, and tuberculosis. We also collected demographics (e.g., age, gender, education), measured deliberation during risk evaluation (i.e., time spent on the screen with the risk information), and assessed participants' subjective confidence in their understanding of the risks (e.g., confidence).

2.1. Design

The main manipulations of the modified mixed-factorial experiment (Figure 2) featured a 2 (outcome label: Ebola hemorrhagic fever, Viral hemorrhagic fever) by 2 design (probability label: risk, likelihood). All participants received the same information about the Ebola virus, however, half of the participants received the information about the disease with its popular name "Ebola hemorrhagic fever" while the other half saw "Viral hemorrhagic fever". The latter represents a broader class of fevers including Ebola. In addition, when referring to the probability of infection or dying, half of the participants saw the word "risk" while the other half saw the word "likelihood". For example, Figure S1 in the appendix shows all the information provided to participants in the "Ebola" + "Risk" condition. The manipulations of the outcome and probability labels were maintained throughout the experiment every time a reference was made to the virus or the probability of infection. In addition, we included a fifth condition where visual aids accompanied the numerical information about the risks (Figure 3). This condition was a duplicate of the "Ebola" + "risk" condition plus the addition of the visual aids.

2.2. Participants and Procedure

The study took place over a three hour period on the 19th of October 2014, 20 days after the CDC confirmed the first case of Ebola diagnosed in the United States. We initially recruited 508 participants on Amazon's Mechanical Turk Web Panel who were current residents of the United States. Mean age was 38 years (SD=13, min.=18, max.=79). One hundred-and-seventy-one (34%) participants were men and 337 (66%) were women. Twelve percent had no more than a high school diploma, 36% had completed up to some college or an associate's degree, 36% had a bachelor's degree, and 16% had a master's degree or higher. Thirteen percent were students or trainees, 65% were employed, 18% were unemployed, and 4% were retired. All participants were randomly assigned to groups and completed the study in about 15 minutes. Three participants who had excessively long completions times were excluded from subsequent analyses resulting in a final sample of 505.

Participant recruiting targeted individuals residing in the U.S. who were invited to participate in a study about health decisions in exchange for $0.55. Recruited participants were asked not to use a calculator or consult any external sources of information during the survey. All participants next completed the adaptive Berlin Numeracy Test for the general population, read information about the virus, and answered a series of questions assessing comprehension, affective reactions, behavioral intentions, and policy endorsements. The study ended with a personality assessment and questions about demographics. Ethical approval for this study was obtained from the Michigan Technological University Review Board. Participants signed the approved electronic informed consent form at the beginning of the study and received the required debriefing upon study completion.

2.3. Materials and dependent measures

Numeracy. Numeracy was measured with the adaptive Berlin Numeracy Test BNT-S version that includes three items from Schwarz et al. (Schwartz et al., 1997) and 2 to 3 items from the adaptive version of the Berlin Numeracy Test (Cokely et al., 2012). For example: "Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips?". Following recommended and validated procedures (Cokely et al., 2012), we computed a sum of participants' weighted correct answers on both tests. The test showed excellent psychometric sensitivity with mean numeracy of the sample at 4.20 (53%), SD=1.65, min.=1, max.=7.

Information about the virus. An example of the information provided to participants is displayed in Figure S1 in the Appendix. Provided information was based on a factsheet on the Ebola virus disease from the World Health Organization (WHO) (World Health Organization, 2014). The factsheet provided data on the probability of dying once infected with the disease (i.e., 50%). Given the emerging nature of infection in the U.S. there were no official estimates of the probability that U.S. residents would be infected with the virus. To provide a plausible numeric estimate of the risk we obtained the number of registered Ebola cases in the U.S. at the date of the survey from the website of the U.S. centers for Disease Control and Prevention (CDC, www.cdc.gov) and divided these by the estimated U.S. population size. These gave us a probability of 1 in 100 million (.000001%) for infection with the virus. The probability of dying among those not infected with the virus was thus 1 in 200 million (.0000005%), consistent with a 50% chance of death after infection as indicated in the WHO factsheet. The small size of the risk accords with currently available simulation data provided by various risk management professional groups (e.g., Risk Management Solutions, www.rms.com).

Risk comprehension. We assessed risk comprehension with two types of items displayed in Table 1. We administered six items designed to assess participants' comprehension of the absolute risks and six items designed to assess participants' comprehension of comparative risks. The second set of items compared the risk of Ebola to other more probable health and safety risks. The twelve items showed good internal consistency (Cronbach's alpha=.79). To arrive at a final risk comprehension score we computed an average of the proportions of correct answers for each set of items. The resulting score had a desirable psychometric profile with a mean=.58, SD=.25, and range from 0 to 1.

Confidence. For each of the absolute risk items, participants were asked to indicate how confident they were that they had given a correct answer on 11-point scales ranging from 0% confident (completely guessing) to 100% confident (completely certain) in intervals of 10%. We computed participants' mean confidence across the five items. The final score exhibited desirable psychometrics with a mean=6.16, SD=2.57, min.=1, max.=11.

Deliberation. We recorded how much time participants spent on the page with risk information before proceeding to the page with the comprehension questions. Typical for reaction time measures, the variable had a positive skew of 11.51 (SE=.11). Three participants were influential outliers, such that they spent unusual amounts of time on the information page (far more than 3 SDs from the mean). These individuals were excluded from further analyses including this variable.

Affective reactions. We assessed affective responses with items developed for health behavior research (Rothman, Martino, Bedell, Detweiler, & Salovey, 1999). The full instrument has been used in published risk communication research since 2011 and is currently in the final stages of validation for cross-cultural risk communication applications (i.e., the Berlin Emotional Responses to Risk scale). The scale asks participants to indicate how they felt about the disease after reading the information. Specifically, on scales from 1 (not at all) to 7 (extremely) they indicated how assured, calm, cheerful, happy, hopeful, relaxed, relieved, anxious, afraid, discouraged, disturbed, sad, troubled, and worried they felt. The order of the adjectives was randomized. We averaged the scores across all negative adjectives as a measure of negative affect (Cronbach's alpha=.93, mean=3.76, SD=1.66, min.=1, max.=7) and all positive adjectives as a measure of positive affect (Cronbach's alpha=.88, mean=2.62, SD=1.31, min.=1, max.=7).

Behavioral intentions. We assessed participants' intentions to adopt behaviors to reduce the risk of the virus with 19 items. Participants were asked to imagine that the news reported one suspected fever case in their city. On scales from 1 (no change) to 5 (extremely likely to avoid) participants indicated how likely they were to avoid any of 13 behaviors (e.g., going to work, using public transportation, sending their children to school, going to their annual doctor's appointment, playing team sports, dancing with strangers, dining out, etc.). In addition, on scales ranging from 1 (strongly disagree) to 7 (strongly agree) participants indicated their agreement with actions aiming at prevention or protection from the virus (3 items, e.g., The government should introduce stricter border control to keep Ebola/Viral fever from spreading.). We also measured general intentions to adopt risk-reducing behavior (3 items, e.g., I intend to take action to protect myself from Ebola/Viral fever.). The 19 items showed excellent internal consistency with Cronbach's alpha=.94. The final composite behavioral intentions score had a mean=60, SD=22, min.=19, max.=107. Larger score indicated behavioral intentions that were more avoidance-oriented and on average would be more likely to be associated with counterproductive personal, social, economic, and health outcomes. In part, we reasoned that the negative consequences of avoidance-oriented behaviors followed from the essentially trivial risk of Ebola given the nature of transmission of the virus. Such extreme avoidance-oriented behaviors also run counter to the public health policies and recommendations of the CDC for U.S. residents.

Policy recommendations. We presented participants with four scenarios presenting different public policy options. Participants were asked to imagine that the U.S. government or the World Health Organization was considering redirecting funds currently used to combat other diseases towards Ebola/Viral fever prevention and treatment, including: (1) funds for cardiovascular disease prevention, (2) funds for development of HIV/AIDS vaccine, (3) funds covering the treatment of 10,000 individuals infected with malaria, and (4) funds covering the treatment of 10,000 individuals infected with tuberculosis. All four cases assessed the degree to which one recommended that policies be changed in favor of prevention or treatment of individuals with Ebola/Viral fever. Participants indicated their support for these policies on scales from 1 (strongly against) to 7 (strongly in favor). The 4 items had good internal consistency with Cronbach's alpha=.89. The final score had a mean=3.14, SD=1.54, min.=1, max.=7, such that a larger score indicated stronger endorsement of these policies; however, some final analyses reverse scoring for ease of interpretation (e.g., higher as better). We chose these diseases because the risk of dying from each disease is much higher than the risk of dying from Ebola. We also chose the number of affected individuals to be larger than the number of registered Ebola cases in the world at that time. Given the assumption that on average better policies should protect and save more lives, we assumed that better recommendation decisions were those that did not redirect funds to combat Ebola/Viral fever.

Table 1. Questions used to assess risk comprehension. Absolute risk items 1 and 2, and 3 and 4 respectively, contained the same question and differed in the response format offered to participants. % correct – percentage of participants who gave a correct answer to the question.

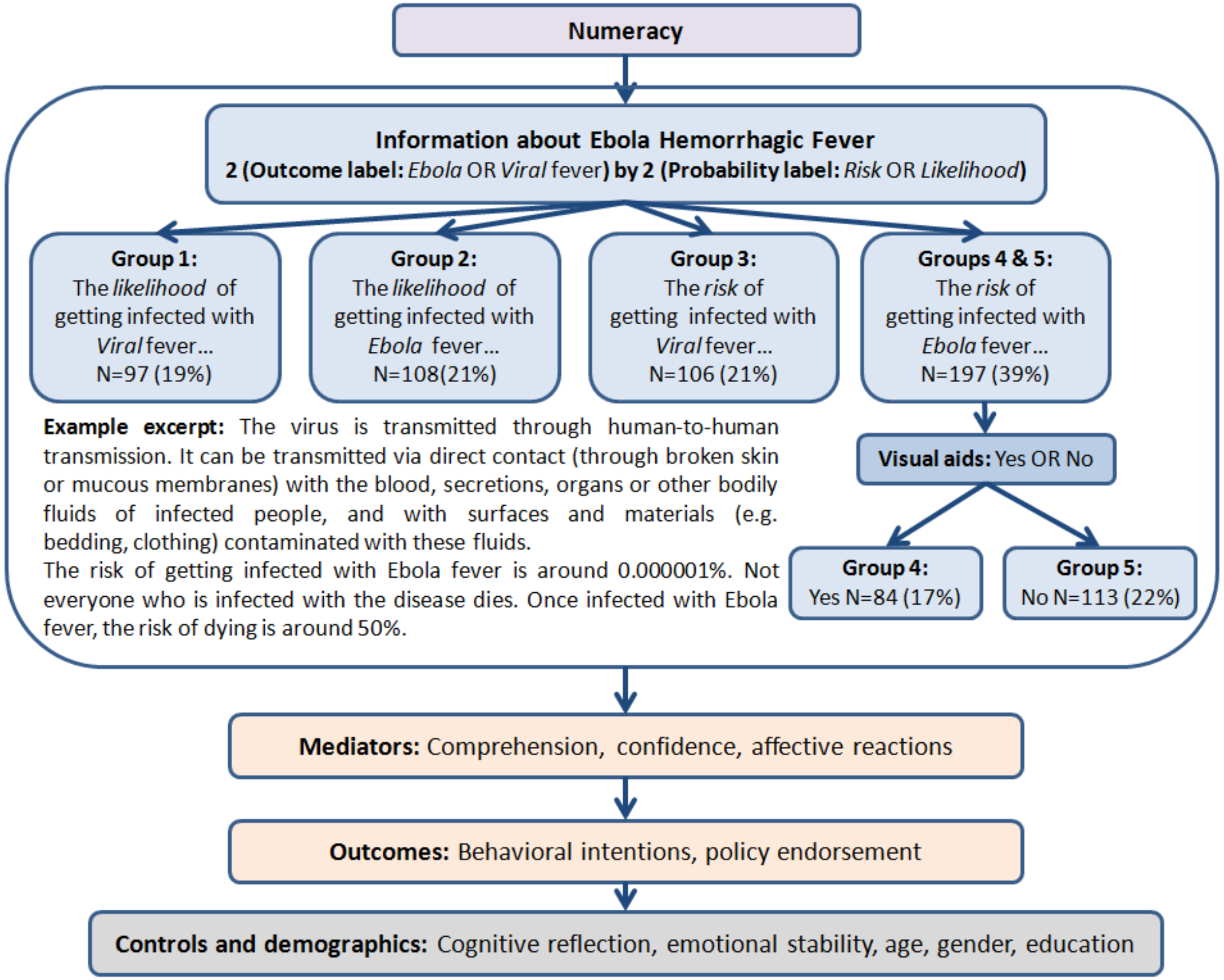
Note: The exact correct answers to absolute risk items 1 and 3 are .02 and .01, respectively. However, because this means less than one person and can be especially difficult for participants, we regarded any value ≤1 as correct.

Absolute risks	**% correct**
Imagine 2,000,000 individuals from an average city. According to the information provided, please give your best estimate of the following: How many people in this average city will be infected with Ebola/Viral fever?	
[item 1] ___ out of 2,000,000 individuals *(Correct: ≤1)*	20
[item 2] A. 1 or fewer; B. around 10; C. around 50; D. around 100; E. around 250; F: around 500; G: around 1,000; H: around 5,000; I. around 10,000; J: more than 10,000 *(Correct A)*	46
Imagine 2,000,000 individuals from an average city. According to the information provided, please give your best estimate of the following: How many people in this average city will die from Ebola/Viral fever?	
[item 3] ___ out of 2,000,000 individuals *(Correct: ≤1)*	51
[item 4] A. 1 or fewer; B. around 10; C. around 50; D. around 100; E. around 250; F: around 500; G: around 1,000; H: around 5,000; I. around 10,000; J: more than 10,000 *(Correct A)*	58
Imagine 300 individuals who have been infected with Ebola/Viral fever. According to the information provided, please give your best estimate of the following: How many of them will die from the disease?	
[item 5] ___ out of 300 individuals *(Correct: 150)*	71
[item 6] *Correct if answer to item3=50% of answer to item 1.*	48

Table 1 (continued).

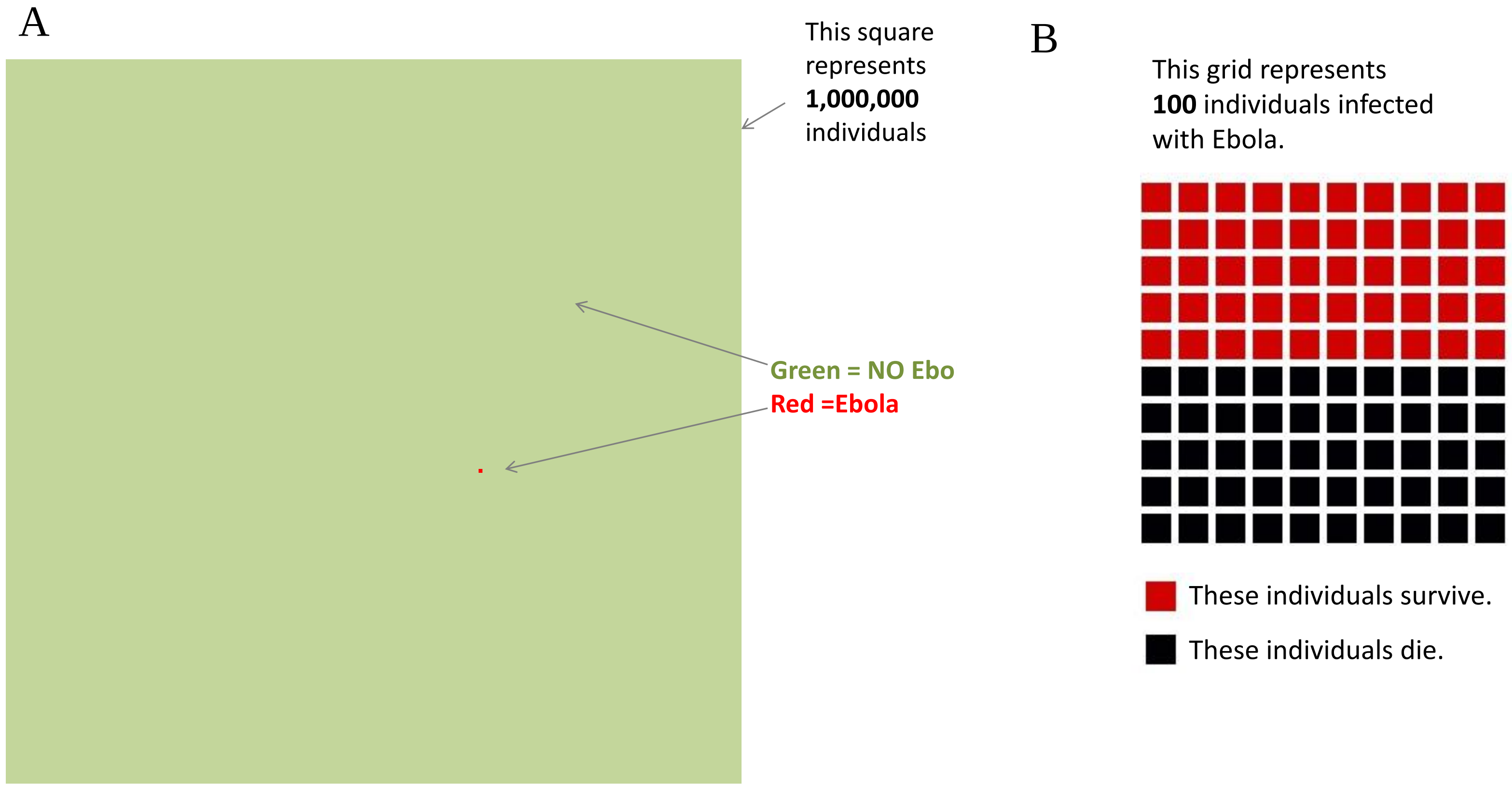
Comparative risks	**% correct**
What is more probable: That an average person dies from Ebola/Viral fever or …[items 1 to 6]? A. much more probable [item 1 to 6]; B. somewhat more probable [item 1 to 6]; C. both are about equally probable; D. somewhat more probable from Ebola/Viral fever; E. much more probable from Ebola/Viral fever. *(Correct: A for items 1 to 6).*	
[item 1]… from heart disease?	82
[item 2]… from cancer?	74
[item 3]… from stroke?	72
[item 4]… in a traffic accident?	78
[item 5]… from assault with a firearm?	62
[item 6]… in an air travel accident?	31

Figure 2. Short description of the experimental design and procedure. The excerpt depicts part of the information that was communicated to the participants.



Example excerpt: The virus is transmitted through human-to-human transmission. It can be transmitted via direct contact (through broken skin or mucous membranes) with the blood, secretions, organs or other bodily fluids of infected people, and with surfaces and materials (e.g. bedding, clothing) contaminated with these fluids.
The risk of getting infected with Ebola fever is around 0.000001%. Not everyone who is infected with the disease dies. Once infected with Ebola fever, the risk of dying is around 50%.

Figure 3. Visual aids used to illustrate the small risk of getting infected with Ebola (A) and the risk of dying once infected (B). Statistics are from the World Health Organization (World Health Organization, 2014).



3. Results

3.1. Main Effect Analyses. To illustrate and contextualize key findings first we discuss significant main effects that were expressed independent of all other assessed influences caused by the experimental manipulations or other psychological variables. Please see next section with structural modeling for precise statistical estimates and methods.

Numeracy and Decision Aids. The benefits of numeracy were considerable. Residents with higher numeracy understood the danger of infection and death from Ebola much better than those with lower numeracy (Figure 4). In turn, observed differences in comprehension that were associated with numeracy translated into reductions in fear, decreases in unhealthy behavioral intentions, and increased rates of policy recommendations likely to save lives. Similarly, the presence of simple visual aids casually improved comprehension across all participants (Figure 5), which also translated into reduced fear, healthier behavioral intentions, and higher rates of life-saving policy recommendations.

Message Framing. Framing risk communications about the disease in terms of "Ebola Fever" versus "Viral Fever" significantly increased negative emotions including fear and distress (Figure 6), which in turn encouraged unhealthy behavioral intentions like canceling annual check-ups, avoiding hospitals despite the need for medical attention, and keeping children home from school (Figure 7). Framing communications in terms of "Ebola" instead of viral fever also reduced support for life-saving policies, prompting participants to recommend the defunding of programs for HIV/AIDs, Maleria, Tuberculosis, and heart disease prevention in favor of funding new Ebola treatment programs in the U.S. (Figure 8). Similarly, framing risk communication information using the term "risk" instead of "likelihood" caused small but significant reductions in risk comprehension across all participants independent of other factors including other experimental manipulations (Figure 9). This finding appears to be the first evidence that the word "risk" itself may cause confusion biasing judgment in risk communications, even among highly numerate individuals. Note that both framing effects biased decisions outcomes (i.e., behavioral intentions and policy recommendations) via influences on different decision processes. While frames emphasizing "Ebola" biased decisions by increasing fear without affecting comprehension, message frames emphasizing "risks" biased decisions by decreasing comprehension without affecting emotional reactions.

Figure 4. The graph represents the level of risk comprehension as a function of numeracy. Low numeracy was identified as the bottom quartile and high numeracy was identified as the top quartile in numeracy scores. The y-axis displays the dependent variable as a proportion of maximum risk comprehension such that a higher score means greater comprehension after covarying the three other conditions (Ebola versus Viral Fever, Visual aid versus No Visual Aids, and 'risk' versus 'likelihood').

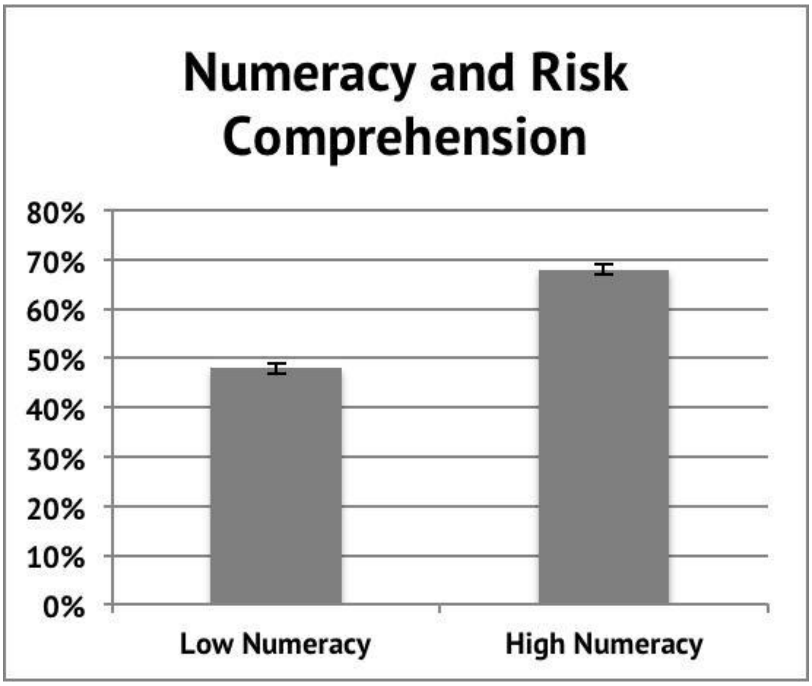


Figure 5. The graph represents the effect of visual aids on risk comprehension. The y-axis displays the dependent variable as a proportion of maximum risk comprehension such that a higher score means greater comprehension after covarying the two other conditions (Ebola versus Viral Fever and 'risk' versus 'likelihood').

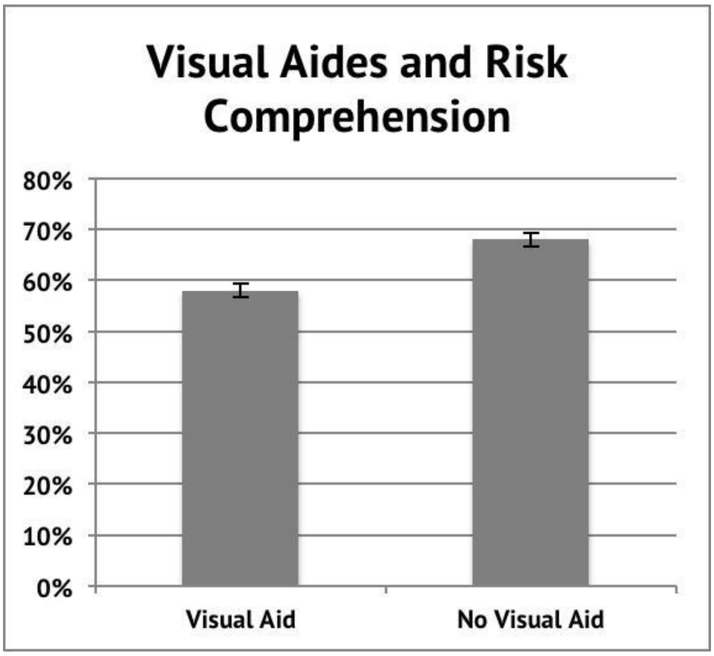


Figure 6. The graph represents the effect of 'Ebola' versus 'Viral fever' on negative emotional reactions. The y-axis displays the dependent variables as a proportion of maximum negative emotional reactions such that a higher score means greater negative emotional reactions after covarying the two other conditions ('Risk' versus 'Likelihood' and Visual aid versus No Visual Aids).

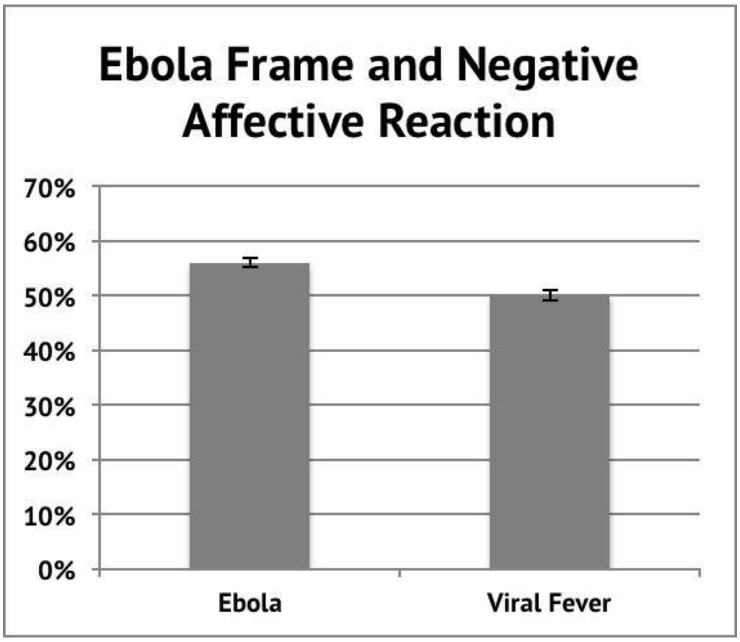


Figure 7. The graph represents the effect of 'Ebola' versus 'Viral fever' on healthy behavioral intentions. The y-axis displays the dependent variables as a proportion of maximum healthy behavioral intentions such that a higher score means healthier behavioral intentions after covarying the two other conditions ('Risk' versus 'Likelihood' and Visual aid versus No Visual Aids).

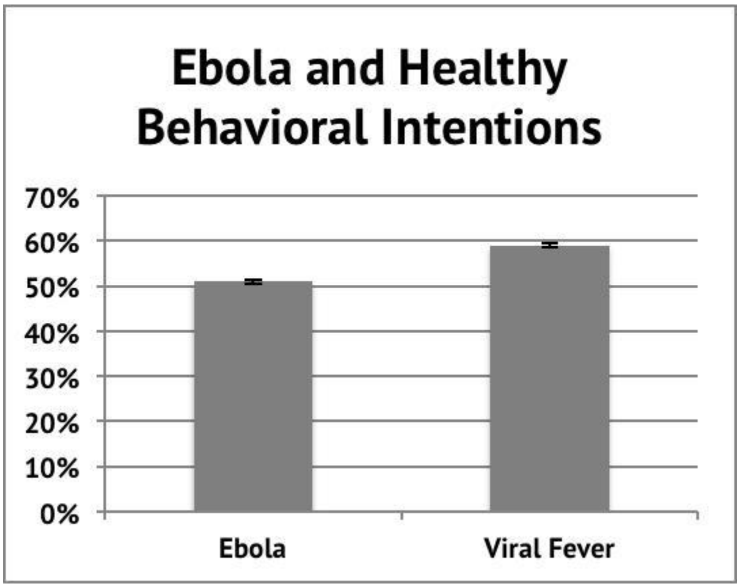


Figure 8. The graph represents the effect of 'Ebola' versus 'Viral fever' on normative policy recommendations. The y-axis displays the dependent variables as a proportion of maximum normative policy recommendation such that a higher score means recommendation of more normatively correct policies after covarying the two other conditions ('Risk' versus 'Likelihood' and Visual aid versus No Visual Aids).

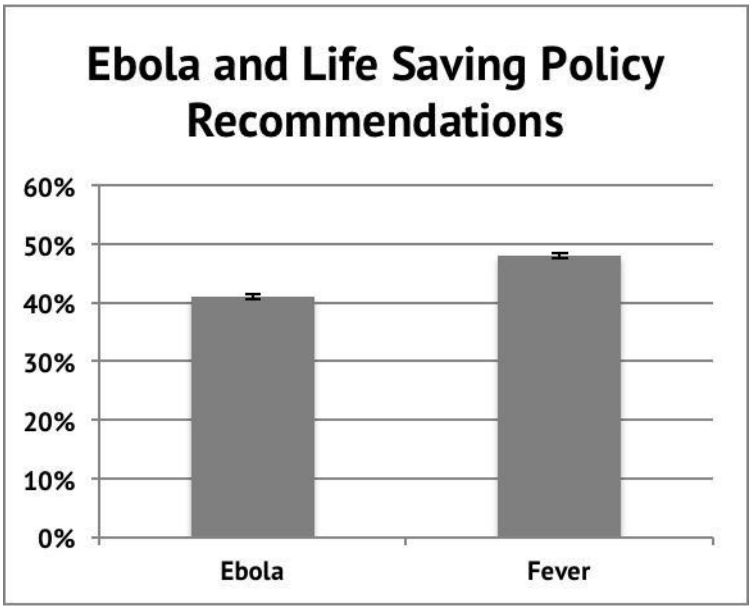
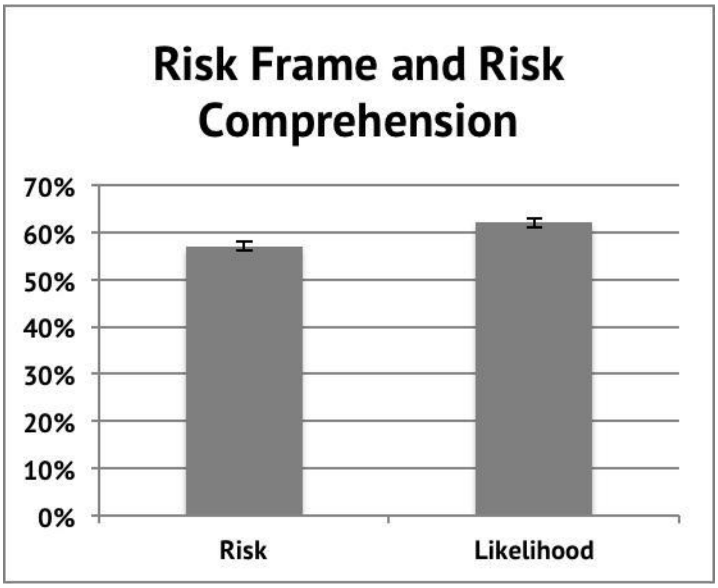


Figure 9. The graph represents the effect of framing the problem in terms of 'risk' versus 'likelihood'. The y-axis displays the dependent variabls as a proportion of maximum risk comprehension such that a higher score means greater comprehension after covarying the two other conditions (Ebola versus Viral Fever and Visual aid versus No Visual Aids).



Structural Decision Process Modeling. Research indicates that the link between numeracy and superior decision making under risk often follows because numerate people tend to use heuristic deliberation strategies to elaborate on risks and consequences, enhancing their ability to assess their own comprehension (e.g., less overconfidence), and fine-tuning affective reactions while improving their understanding of the decision gist (Cokely et al., 2012; Garcia-Retamero & Cokely, 2013). To evaluate this theoretical framework, we used structural equation modeling (SEM) to construct and competitively test specific decision process models mapping direct and indirect effects of numeracy, risk labeling, visual aids, and Ebola labeling on (a) Behavioral Intentions and (b) Policy recommendation decisions. We hypothesized that the effects of these variables would be mediated by time spent deliberating, subjective confidence in one's own risk comprehension, actual risk comprehension, and the strength of negative affective reactions (see a priori specified models in Figures 10 and 11).

SEM techniques were used to estimate model fits separately for the observed data for (a) Behavioral Intentions and (b) Policy Decision outcomes. The two isomorphic structures of our theory-based models are presented in Figures 10 and 11. Testing of the theory-based models began by evaluating consistency with the observed data. Analyses indicated that both models robustly explained the observed data passing all conventional fit criteria irrespective of considerable model complexity (see Tables 2 and 3) (Kline, 2011). Chi-squared tests revealed no significant deviations or model fit violations. RMSEAs estimates were less than .06 and the 90% confidence intervals around the RMSEAs included 0. The *pclose* test was non-significant. The CFIs and TFIs were greater than .95, and the SRMRs were less than .08. Overall, the full models were strong predictors of decision outcomes variables, explaining between 23-34% of the total observed variance.

Although our theory-based models predicted the data well, there could be other models that fit the observed data better. In accord with best practices for modeling, and to help test whether our theory-based models offered the most efficient theoretical account of the data, six alternative models for the data were specified testing theoretically interesting plausible alternatives (i.e., three for Behavioral Intentions and three for Policy recommendations). For example, we tested whether model fit improved if we assumed that affect gave rise to better comprehension (e.g., people who were less fearful might be more willing to think more carefully about risks) or if we assumed that comprehension preceded deliberation (e.g., people who understood risks better might be more likely to deliberate and become confident in their own understanding). For each of the three models we re-ordered the sequence of the 4 endogenous variables (Deliberation, Confidence, Comprehension, and Affect) but left all other relations unchanged. Specifically, we used the following alternative sequences of endogenous variables: *Model 2* Affect→ Deliberation → Confidence → Comprehension; *Model 3* Comprehension

→ Deliberation → Confidence → Affect; and *Model 4* Deliberation → Comprehension → Confidence → Affect. *Model 1* represents the theory-based *a priori* model. Separate path analyses were conducted and evaluated for each of the alternative models.

Results of the competitive structural decision process modeling suggested that our theory-based models (i.e., Model 1) provided a significantly and substantially better account of observed data than any of the alternative models. For Policy Decisions, Models 3 and 4 failed the chi-squared test suggesting significant misfit of the hypothesized models and the observed data. While Model 2 passed the chi-squared test, Model 1 had lower AIC and BIC values indicating that the theory-based model provided more consistent and robust explanatory power. A similar pattern emerged for Behavioral Intentions, where Model 2 failed the chi-squared test. Although models 3 and 4 passed the chi-squared test, the AIC and BIC values were lower for Model 1 than for either Model 3 or 4. On the whole, results suggest that for theoretical and empirical reasons our *a priori* theory-based Model 1 represents the most-comprehensive, robust, and parsimonious explanation of the observed data as compared to other theoretically plausible alternative models.

Direct, indirect, and total effects of the theory-based process models are reported in the Appendix. All anticipated direct effects were found and the series of predicted indirect effects was also documented. While there is no simple standardly accepted way to assess indirect effects with multiple-mediators in structural equation models (Kline, 2011), one rule of thumb is that if a series of path coefficients are significantly related then the entire indirect effect path is significant as well (Kline, 2011; Cohen & Cohen, 1983). In this case, since numeracy follows a path with all significant path coefficients leading to the outcome variables (i.e., Behavioral Intention or Policy), the entire indirect effect can be assumed to be significant.

Accordingly, structural process modeling indicates that more numerate people deliberated more, leading those people to have more well-founded confidence in their comprehension, improving their understanding, reducing affective reactivity, and ultimately leading to better policy decisions and healthier behavioral intentions. Numeracy also had direct effects on the 4 mediating variables, indicating that more numerate people tended also to be more confident, understand more, and have less extreme affective reactions, independent of other indirect effects.

While numeracy is featured prominently in the two theory-based structural decision process models, visual aids were also found to causally contribute to better policy decisions and behavioral intentions through essentially the same decision process pathways. Specifically, the presence of visual aids increased confidence and comprehension, thereby attenuating negative affective reactions and resulting in better overall decisions (see Appendix). There is good reason to think that the reason decision aids did not also have direct and indirect effects through deliberation is because the availability of the visual aids changed the amount of time needed to evaluate risks (e.g., a picture can be worth a thousand words).

Figure 9. SEM decision process model with standardized path coefficients and error terms for the Behavioral Intention outcome variable. N=505. The appendix reports the direct, indirect, and total effects in Figure 9.

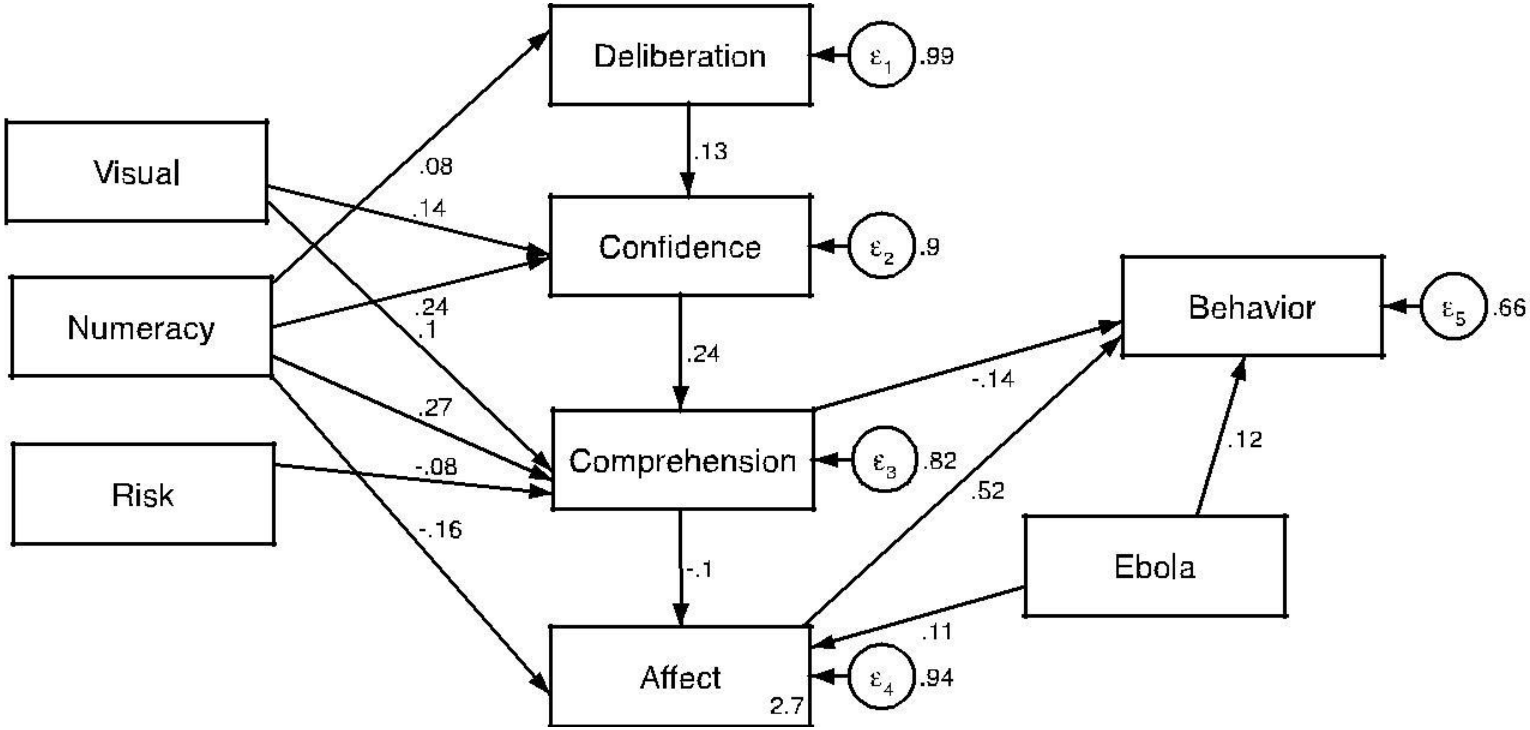


Figure 10. SEM decision process model with standardized path coefficients and error terms for the Policy outcome variable. N = 505. The appendix reports the direct, indirect, and total effects from Figure 10.

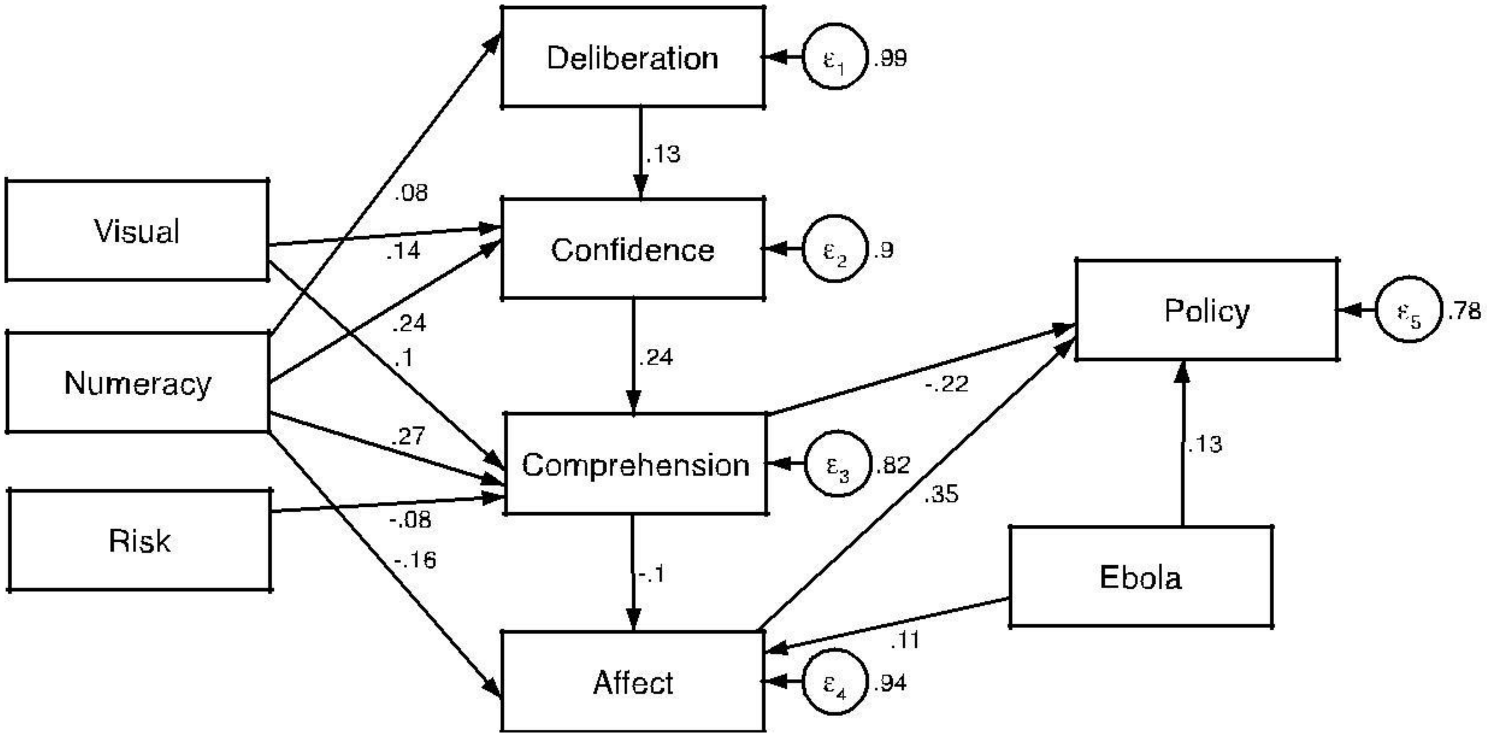


Table 2. Model comparisons for Behavioral Intentions. Model 1 is the a priori preferred model in Figure 9. Model 2 is modified so that Affect → Deliberation → Confidence → Comprehension. Model 3 is modified so that Comprehension → Deliberation → Confidence → Affect. Model 3 is modified so that Deliberation → Comprehension → Confidence → Affect.

Model	*df*	*Chi*^{2}	*p*	RMSEA	90% CI	Pclose	AIC	BIC	CFI	TLI	SRMR
1	16	11.43	.78	0	0-.03	.99	18876	18978	1	1	.02
2	16	15.8	.47	0	0-.04	.99	18881	18982	1	1	.02
3	16	44.49	< .001	.06	.04-.08	.21	18909	19011	.92	.86	.04
4	17	20.55	.25	.02	0-.05	.97	18884	18981	.99	.98	.02

Table 3. Model comparisons for Policy. Model 1 is the a priori preferred model in the in Figure 10. Model 2 is modified so that Affect → Deliberation → Confidence → Comprehension. Model 3 is modified so that Comprehension → Deliberation → Confidence → Affect. Model 3 is modified so that Deliberation → Comprehension → Confidence → Affect.

Model	*df*	*Chi*^{2}	*p*	RMSEA	90% CI	Pclose	AIC	BIC	CFI	TLI	SRMR
1	16	11.91	.75	0	0-.03	.998	16285	16387	1	1	.018
2	16	16.29	.43	.01	0-.04	.98	16290	16391	.99	.99	.02
3	16	44.97	<.01	.06	.04-.08	.2	16318	16420	.9	.81	.04
4	17	47.21	<.01	.06	.04-.08	.2	16319	16416	.9	.82	.04

4. Discussion

When faced with an unprecedented emerging national crisis that poses substantial threats to personal, political, and economic stability, should we tell people the truth or do we need to protect people from their own irrational over-reactions and biases? When strong emotions and conflicts of interest in the media prompt wild and provocative speculation, can we trust people to see through the noise and make well-informed decisions on the basis of simple and honest risk communications? To address these and other questions, three weeks after the first confirmed case of Ebola infection in the U.S., at the peak of the worst Ebola pandemic in history we conducted a risk communication experiment testing a structural process model of informed decision making under dread risk. Analyses including competitive structural model testing revealed strong and central influences of risk comprehension on informed decision making, despite opposing biases (e.g., affective responses & message framing). Consistent with general models of decision making under risk, numeracy was found to be a strong and consistent predictor of adaptive decision making because it directly and indirectly influenced risk comprehension and affective reactions via increased deliberation and confidence (i.e., metacognition, Cokely & Kelley, 2009; Cokely et al., 2012; Garcia-Retamero et al., 2015). Although practical mathematical skills are among the most influential educational factors contributing to economic prosperity in industrialized countries (Hunt & Wittmann, 2008), the current results suggest that the benefits of a well-informed and numerate populous do not essentially STEM from the increased availability of well-trained professionals in science, technology, engineering, and mathematical fields. In contrast, even small differences in numeracy may have profound effects on risk comprehension and everyday choices, reducing costly personal, social, and economic inefficiencies (Cokely et al., 2012; Peters et al., 2007; Peters et al., 2010; Reyna et al., 2009).

The observed effects of visual aids in the current study provide another compelling demonstration that less numerate adults can and often do make good decisions when information formats help them understand risks (Bruine de Bruin et al., 2007; Bruine de Bruin & Bostrom, 2013; Fischhoff, Brewer, & Downs, 2012; Fischhoff, 2013; Garcia-Retamero & Cokely, 2013; Trevena et al., 2013). Contrary to some popular appeals, well-informed decisions about complex, high-stakes issues do not typically require uncommon levels of intelligence or special abstract reasoning capacities. Instead, numeracy and risk literacy promote adaptive decision making by (i) encouraging deliberate thinking about thinking (i.e., metacognitive heuristics), (ii) promoting meaning-oriented comprehension (e.g., decision gist), and (iii) fine-tuning affective reactions. In turn, one's evaluation and understanding of risk promotes more adaptive decision making because it calibrates simple heuristics allowing decision makers to identify and feel the weight of the most essential information, while ignoring irrelevant, redundant, and non-diagnostic cues (Cokely et al., 2012; Gigerenzer et al., 1999; Peters, 2012; Petrova, van der Pligt, & Garcia-Retamero, 2014; Reyna et al., 2009). These results accord with many others indicating that risk literate decision makers should not aspire to become cold and calculating logical optimizers who just "do the math". In our fundamentally uncertain and complex world, simple heuristic decision strategies can outperform even the most complex optimization processes (Gigerenzer et al., 1999; Gigerener & Gaissmaier, 2011).

Beyond modeling decision mechanisms and consequences, the current results speak to fundamental questions about informed decision making in a modern mixed economy democracy. There is now a great debate underway concerning advances in decision psychology that use paternalistic mechanisms like persuasion and choice architecture to profitably nudge decision making one way or another. Although government sponsored liberally paternalistic policies are intended to promote people's own best interests while preserving choices, non-rational persuasion techniques usually influence behavior without explicit consent or meaningful comprehension on the part of decision makers (e.g., people often just accept defaults set by trusted sources for better and worse). To be clear, the great debate is not about whether or not nudging is "better" or "worse" than informed decision making. Instead, the heart of the debate concerns "when" and "why" nudging or informed decision making should be preferred, which is an ethical and empirical question that psychological science is uniquely positioned to address. Ultimately, paternalistic policies tend to infringe on autonomy and thus depend heavily on political agendas and the wisdom of current policy-makers and institutions (Gigerenzer 2015). In contrast, investments that enhance risk literacy like education and policies that promote honest risk communications help individuals make decisions on their own in justifiable and adaptive ways, promoting shared and informed participation in decision making and governance (Bruine de Bruin & Bostrom, 2013; Gigerenzer, 2014). Although it seems likely that some risks will only be manageable with bold and persuasive action by administrators, the current results add to a growing body of research indicating that even when the stakes are high and emotions are running hot, factors that promote risk literacy empower adaptive and informed decision making.

5. References

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—harnessing the Web for public health surveillance. *New England Journal of Medicine*, *360*(21), 2153-2157.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology, 92*(5), 938-956.

Bruine de Bruin, W., & Bostrom, A. (2013). Assessing what to address in science communication. *Proceedings of the National Academy of Sciences of the United States of America, 110 Suppl 3*, 14062-14068.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*, 25-47.

Cokely, E.T., Ghazal, S., Galesic, M., Garcia-Retamero, R., & Schulz, E. (2013). How to measure risk comprehension in educated samples. In R. Garcia-Retamero & M. Galesic (Ed.), *Transparent communication of health risks: Overcoming cultural differences* (pp. 29-52). New York: Springer.

Drane, J. F. (1984). Competency to give an informed consent: A model for making clinical assessments. *JAMA, 252*(7), 925-927.

Fischhoff, B., Brewer, N. T., & Downs, J. S. (2012). *Communicating risks and benefits: An evidence based user's guide.* Government Printing Office.

Fischhoff, B. (2013). The sciences of science communication. *Proceedings of the National Academy of Sciences of the United States of America, 110 Suppl 3*, 14033-14039.

Gaissmaier, W., & Gigerenzer, G. (2012). 9/11, act II: A fine-grained analysis of regional variations in traffic fatalities in the aftermath of the terrorist attacks. *Psychological Science, 23*(12), 1449-1454.

Garcia-Retamero, R., & Galesic, M. (2010). Who proficts from visual aids: Overcoming challenges in people's understanding of risks. *Social Science & Medicine, 70*(7), 1019-1025.

Garcia-Retamero, R., & Galesic, M. (2013). *Transparent communication of health risks: Overcoming cultural differences*. New York: Springer.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22*(5), 392-399.

Garcia-Retamero, R., & Cokely, E. T. (2014). The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making, 27*(2), 179-189.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Gigerenzer, G. (2006). Out of the frying pan into the fire: Behavioral reactions to terrorist attacks. *Risk Analysis, 26*(2), 347-351.

Gigerenzer, G. (2014). *Risk savvy: How to make good decisions.* Penguin.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684-704.

Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012-1014.

Hayes, A. F. (2008). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Levin, I. P., Bossard, E. A., Gaeth, G. J., & Yan, H. (2014). The combined role of task, child's age and individual differences in understanding decision processes. *Judgment and Decision Making, 9*(3), 274-286.

Kline, R. B. (2011). Convergence of structural equation modeling and multilevel modeling. In Williams, M., & Vogt, W. P. (Eds.), *The SAGE Handbook of Innovation in Social Research Methods* (pp. 562-589). Sage Publications.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*(1), 37-44.

Mellon, J. (2013). Where and when can we use Google Trends to measure issue salience?. *Political Science & Politics*, *46*(02), 280-290.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine*, *35*(3), 261-274.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making, 25*(4), 390-401.

Parker, A. M., & Fischhoff, B. (2005). Decision- making competence: External validation through an individual- differences approach. *Journal of Behavioral Decision Making, 18*(1), 1-27.

Peters, E., Baker, D. P., Dieckmann, N. F., Leon, J., & Collins, J. (2010). Explaining the effect of education on health: A field study in Ghana. *Psychological Science, 21*(10), 1369-1376.

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs (Project Hope), 26*(3), 741-748.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31-35.

Petrova, D. G., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making, 27*, 191-199.

Reyna, V. F. (2004). How people make decisions that involve risk A dual-processes approach. *Current Directions in Psychological Science, 13*(2), 60-66.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943-973.

Ripberger, J. T. (2011). Capturing curiosity: Using Internet search trends to measure public attentiveness. *Policy Studies Journal*, *39*(2), 239-259.

Rothman, A. J., Martino, S. C., Bedell, B. T., Detweiler, J. B., & Salovey, P. (1999). The systematic influence of gain-and loss-framed messages on interest in and use of different types of health behavior. *Personality and Social Psychology Bulletin, 25*(11), 1355-1369.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*(11), 966-972.

Slovic, P. (1987). Perception of risk. *Science, 236*(4799), 280-285.

Swearingen, C. D., & Ripberger, J. T. (2014). Google Insights and US senate elections: does search traffic provide a valid measure of public attention to political Candidates?. *Social Science Quarterly*, *95*(3), 882-893.

World Health Organization. (2014.). *Ebola virus disease.* Fact sheet N°103.

CHAPTER 5:
Self-Other Discrepancies in Decision Making

The content of this chapter will be published as Petrova, D., Garcia-Retamero, R., & van der Pligt, J. (in press). What factors moderate self-other discrepancies in decisions? Results from a vaccination scenario. *The Spanish Journal of Psychology*. The appendix is found in this link: https://db.tt/NkvRmX2h.

Chapter 5

What Factors Moderate Self-Other Discrepancies in Decision Making?

When we make risky decisions for others, we tend to follow social norms about risks. This often results in making different decisions for others than we would make for ourselves in a similar situation (i.e., self-other discrepancies). In an experiment, we investigated self-other discrepancies in young adults' decisions to purchase a vaccine against a sexually-transmitted virus for themselves or for another person (i.e., the target of the decision). When the target's preferences were in line with social norms, surrogates showed large self-other discrepancies in line with these norms. When the target's preferences were contrary to social norms, surrogates did not show self-other discrepancies in line with these preferences; instead they still followed social norms Surrogates with lower numeracy and higher empathy showed self-other discrepancies more in line with the target's preferences, even when these were contrary to the norm. Surrogates whose own risk attitudes were contrary to social norms showed larger self-other discrepancies. These results demonstrate that perceived social norms about risk can predict self-other discrepancies in risky decisions, even when the target's preferences are known and at odds with the social norm. Further, the surrogates' numeracy, empathy, and propensity to take risks influence the extent to which risky decisions for others resemble risky decisions for oneself.

1. Introduction

People often make risky decisions for others in their capacity as professionals, significant others, or friends. For example, in the context of health and medicine, doctors frequently make risky treatment decisions for their patients, and family members often decide about prognostic risks that pertain to the well-being of relatives. Research has shown that when making decisions for other people, we often use our own preferences as an anchor (Epley, Keysar, Van Boven, & Gilovich, 2004; Marks & Arkes, 2008). To illustrate, surrogates' predictions of patients' decisions more strongly resemble surrogate's own preferences rather than preferences of patients (Fagerlin, Ditto, Danks, & Houts, 2001). Consistent with this finding, several studies have documented important mispredictions and self-other discrepancies between the wishes of patients and decisions of next of kin surrogates (Shalowitz, Garrett-Mayer, & Wendler, 2006), doctors (Garcia-Retamero & Galesic, 2012, 2014), and parents (Brody, Annett, Scherer, Perryman, & Cofrin, 2005). Knowledge of what factors influence decisions for oneself and others can help facilitate the decision making process and achieve optimal outcomes. Surprisingly, empirical work on how people make decisions for others in comparison to decisions for themselves is relatively scant (Stone & Allgaier, 2008). We extend this research by investigating the differences in risky decisions made for oneself and another person (i.e., self-other discrepancies), depending on the preferences of the other person and the characteristics of the surrogate.

Two empirically-supported theories have informed our research. The first is the theory of the *empathy gap* (or *risk-as-feelings*, Hsee & Weber, 1997; Loewenstein, 2005), mainly based on research investigating *predictions of others' decisions*. This research showedthat people often expect others to have relatively muted emotional reactions toward risks, resulting in regressive predictions of others' decisions (Faro & Rottenstreich, 2006; Loewenstein, 2005). To illustrate, individuals who predicted that others would experience less worry than themselves when faced with a risky health decision (i.e., showed an empathy gap) also predicted that others would make less risk-averse decisions than themselves (i.e., made a regressive prediction) (Garcia- Retamero, Okan, & Maldonado, 2015). More support for the empathy gap comes from research showing that when empathy is facilitated, self-other discrepancies disappear (Hsee& Weber, 1997). Self–other discrepancies in predictions of decisions primarily occurred when the target of the prediction was unfamiliar and/or abstract (i.e., when the other is unknown), but not when the target was vivid (i.e., a person sitting next to us; Hsee & Weber, 1997) or familiar (i.e., a close friend; Faro & Rottenstreich, 2006).

Another theory is *social values theory*, which is based on research investigating whether people are more or less risk-averse in *decisions* for others than for themselves (Beisswanger, Stone, Hupp, & Allgaier, 2003; Dore, Stone, & Buchanan, 2014; Stone & Allgaier, 2008; Stone, Choi, de Bruin, & Mandel, 2013; Wray & Stone, 2005). This research showed that people decide for others in accordance with the perceived social value of the risk in a given domain (Stone & Allgaier, 2008; Stone et al., 2013). In domains where risk-taking is valued (e.g., casual romantic relationships), people make more risk-seeking decisions for others than for themselves. However, in domains where risk-taking is not valued (e.g., situations involving health and safety), people make more risk-averse decisions for others than for themselves. These findings are in accordance with research showing that decision makers consider various aspects of the decision situation when they make decisions for themselves but tend to focus on the most important aspect (e.g., the social norm) when they make decisions for others (Kray, 2000; Kray & Gonzalez, 1999). Finally, the above mentioned self-other discrepancies are not due to a failure to predict what others would decide but rather reflect what is perceived as the normative behavior in a given context (Stone et al., 2013).

These findings show that when we do not know the wishes of other people, we may use the social norm as a cue to make decisions for them. However, on many occasions we know what the other person would prefer because they explicitly told us or we know them well. In addition, the other person's preferences may or may not coincide with the social value of risk in a given domain. It is not yet clear to what extent social values theory can predict decisions when preferences and norms are at odds. To the best of our knowledge, no research has investigated how information about the target's (pro- or contra-normative) preferences affects self-other discrepancies in decisions. In this research we aimed to fill this gap in the literature and expand research on social values theory. In particular, we investigated how information about preferences, and several other cognitive and emotional factors moderated self-other discrepancies in a health decision context.

In an experiment, young adults made decisions about health—a context where risk-aversion and safety are valued (Stone et al., 2013). In particular, participants decided about purchasing a vaccine against a sexually-transmitted virus for themselves and for another person. We manipulated the information participants received about the vaccination preferences of the other person. We expected this information to influence participants' decisions for the other person, thereby producing self-other discrepancies. In one condition, participants received information that the other person had risk-averse preferences (e.g., would rather vaccinate). In a second condition, they received information that the other person had risk-seeking preferences (e.g., would rather avoid the hassle of vaccination). In a third condition, participants received no information.Social values theory predicts that people would make more risk-averse decisions (e.g., would be willing to pay more for vaccination) for the other person than for themselves when they have no information. In contrast, participants may be more risk seeking when they make decisions for others than for themselves in the risk-seeking condition, suggesting that the predictive power of the social norm would be diminished when people's preferences are known to be at odds with the social norm (H_{1a}). Alternatively, if self-other discrepancies are not reversed in the risk-seeking condition, this would mean that under certain conditions the social value of risk is pervasively predictive of self-other discrepancies even when preferences are at odds with the social norm (H_{1b}). In the current research, we tested these two alternative hypotheses.

There is also a lack of research investigating the role of cognitive and emotional individual differences, which could potentially moderate self-other discrepancies. In the current research, we examined the role of individual differences in numeracy, empathic concern, and domain-specific risk taking. Numeracy is the ability to understand and use numerical expressions of probability and has been shown to affect decision making about health risks (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Galesic & Garcia-Retamero, 2011; Peters, 2012). For example, compared to people with high numeracy, people with low numeracy are more influenced by general factors like mood (Västfjäll, Peters, & Starmer, 2011), the credibility of the story narrative (Dieckmann, Slovic, & Peters, 2009), or how the options are framed (Garcia-Retamero & Cokely, 2013; Garcia-Retamero & Galesic, 2010, 2011). Individuals with higher numeracy, on the other hand, deliberate longer (Ghazal, Cokely, & Garcia-Retamero, 2014), and show more precise affective reactions to risks (Peters, 2012; Petrova, van der Pligt, & Garcia-Retamero, 2014). This suggests that, in decisions about health risks, people with high numeracy may show smaller self-other discrepancies by focusing their judgments on objective factors like probabilities, thereby diminishing differences between decisions for themselves and others. People with lower numeracy, on the other hand, may use more general information like the preferences of others to guide decisions for them. Lower numeracy may thus be related to more substantial self-other discrepancies by influencing decisions for others (*H_2*).

We also measured participants' tendency for empathic concern or more specifically "other-oriented" feelings of sympathy and concern for unfortunate others (Davis, 1980). Previous research has shown that greater tendencies towards feeling empathy for others is associated with smaller self-other discrepancies in *predictions* of decisions (Faro & Rottenstreich, 2006; Garcia- Retamero et al., 2015). However, it is not yet clear to what extent empathic concern determines discrepancies between actual decisions for oneself and for others, especially when the wishes of the others are known. People higher in empathy may be more likely to consider the preferences of others when making a decision for them, while people lower in empathy may be less likely to do so. This could results in larger self-other discrepancies for people high in empathy (*H_3*).

Finally, we recorded participants' risk taking propensity in the domain of health (Blais & Weber, 2006). Social value theory posits that people will decide for others based on the perceived social value of risk. This implies that when perceptions of people's own risk behavior match the perceived social value of risk, self-other discrepancies may not exist (i.e., people will decide for themselves and for the other person according to what they perceive to be the social norm). However, the more people's own attitudes and decisions are different from the perceived social norms, the larger self-other discrepancies would be. In this research, we tested to what extent one's own propensity towards risk taking affects self-other discrepancies. Greater risk taking propensity may be related to greater self-other discrepancies by influencing decisions for onself. In particular, when participants are themselves risk-averse, they will make a risk-averse decision for themselves and a risk-averse decision for the other, resulting in little discrepancy between self and other decisions. However, when participants themselves are risk takers, they will make a risk seekingdecision for themselves and a risk-averse decision for the other (consistent with social norms), resulting in larger discrepancy between self andother decisions (*H_4*).

In sum, in an experiment we tested the moderating role of information about the preferences of the target, as well as numeracy, empathy, and risk taking propensity of the decision maker on self-other discrepancies in decisions about health. We measured participants' decisions for themselves and for another person. In particular, we asked participants, given various risks of contracting a virus, how likely it was that they opted for vaccination and how much they would be willing to pay for a vaccine. We also asked participants to what extent they thought they would be worried (when making decisions for themselves) and to what extent the other person would be worried (when they made decisions for the other). Worry is a strong predictor of vaccination decisions (Chapman & Coups, 2006) and can play an important role in self-other discrepancies (Faro & Rottenstreich, 2006; Garcia- Retamero et al., 2015; Loewenstein, 2005). Our aim was to test under what conditions self-other discrepancies in decisions were informed by predicted feelings of worry (e.g., I predict that the other person would be less worried than I would be, so I make a more risky decision for her than I would do for myself, consistent with her *preferences*), or were incongruent with predicted feelings of worry (e.g., I predict that the other person would be less worried than I would be, but despite that I make a more risk-averse decision for her as I would do for myself, consistent with *social values theory*).

2. Method

2.1. Participants

Participants were 144 young adults (mean age = 21, *SD* = 5, 17% male) who completed an online survey in return for course credit or 7€. The experiment was part of a larger online research session administered by the research participation platform of the University of Amsterdam, the Netherlands in 2012.

2.2. Design

The experiment employed a mixed 2 (decision target) by 7 (probability) by 3 (information type) by 2 (order of targets) design. Decision target and probability were manipulated within-subjects; information type and order were manipulated between-subjects.

Participants decided whether they would purchase a vaccine against a sexually-transmitted virus after receiving information about the risk of contracting the virus. They made these decisions for themselves or for a dependent significant other (i.e., a younger sister they were responsible for). The order in which they made decisions for themselves and for their sister was randomized. Participants completed some unrelated filler questions after the first set of decisions.

For every decision target (themselves or their sister), participants read seven analogous scenarios describing different probabilities of contracting the virus. These probabilities were presented as frequencies (i.e., 1, 7, 20, 50, 80, 93, and 99 out of 100 people would get the virus if they do not get vaccinated, respectively) in a semi-randomized order. In particular, participants first received the 1 and 99 out 100 scenarios (order randomized), followed by the remaining scenarios, which were presented in a random order.

In addition, participants were randomly assigned to one of three information type conditions. These conditions differed in the information that participants received about their sister's risk attitude and her vaccination preferences. In the *risk-averse* condition, the sister was described as a person interested in health and someone who would most likely participate in a vaccination program. In the *risk-seeking* condition, the sister was described as a person not interested in health and someone who would rather avoid the hassle of vaccination even if that would entail a somewhat higher risk later in life. Finally, in the *no information* condition, participants did not receive information about their sister's attitudes toward health or the vaccine. A more detailed description of the materials and measures is provided in the online appendix.

2.3. Measures

2.3.1. Self-other discrepancies measures

For each decision target and probability, participants answered three questions. On scales ranging from 1 (not at all) to 100 (very much), participants indicated (1) how likely it was that they would get vaccinated/have their sister vaccinated, and (2) how worried they/their sister would be about contracting the virus. The order of these questions was randomized. Finally, participants also estimated (3) how much they would be willing to pay (WTP) for their/their sister's vaccination. They provided an amount in Euros.We used this measure of WTP as a measure of risk-aversion with potentially high sensitivity (i.e., the more a participant was willing to pay, the more risk-averse).

2.3.2. Individual differences

Numeracy. We measured participants' numeracy with the adaptive version of the Berlin Numeracy Test (Cokely et al., 2012). The test consists of four math-type questions about risks and probabilities and is among the strongest predictors of risk literacy (see RiskLiteracy.org for examples). The test showed good discriminability in this sample with a mean of 2.70 (*SD* = 1.07).

Empathy. We measured empathy with the empathic concern subscale from the Interpersonal Reactivity Index (Davis, 1980). On scales from 1 (absolutely disagree) to 7 (absolutely agree) participants indicated to what extent each of 7 statements described them (e.g., "I often have tender, concerned feelings for people less fortunate than me."). The items showed acceptable internal consistency (Cronbach's alpha = .67). The scale had a mean of 3.74 (*SD* = .54), where a higher score indicated more empathic concern for others.

Risk taking in health. We measured participants' propensity towards risk in the domain of health with two subscales from the Domain-Specific Risk Taking Scale (DOSPERT). In particular, we administered the health and safety (Blais & Weber, 2006) and medical subscales (Butler et al., 2012), each consisting of six items. On scales from 1 (very unlikely) to 7 (very likely) participants indicated to what extent they were likely to perform a certain activity if chance presented itself (e.g., donate blood, have unprotected sex). The health and safety subscale showed acceptable internal consistency with Cronbach's alpha = .66. The medical subscale, however, had a Cronbach's alpha of .52. An examination of the item-total correlations showed that no particular item was responsible for the poor internal consistency. To deal with the low internal consistency and obtain one composite measure of risk taking in the domain of health we combined all 12 items in one scale (Cronbach's alpha = .61) with a mean of 3.87 (*SD* = .79).

2.4. Procedure

Participants were asked to imagine that infections with a new sexually transmitted virus were detected in Europe. We chose this topic because it is a relevant problem for young adults. To make our experiment ecologically valid, the description of the virus was largely based on the human papillomavirus. In order to avoid that participants' previous knowledge and attitudes affected their decisions, the name of the virus was omitted. The virus was described as mostly harmless, but it could also cause cancer with unspecified probability. Participants were told that using condoms or other methods of protection could not provide complete protection against the virus. They were also informed about the recent discovery of a vaccine against the virus. Unfortunately, recent budget cuts in health care precluded a vaccination program paid by the Ministry of Health, and insurance companies also decided not to cover the cost of the vaccine. Thus, the vaccination program was voluntary and participants would have to pay for the vaccine themselves.

Participants received additional information when the target of the decisions was their sister. In particular, they had to imagine that their parents had moved to another country, while they and their sister stayed in Holland to finish their studies. Participants were told that their sister was 17 years old and therefore they were legally responsible for her. Hence, they had to decide whether their sister would participate in the vaccination program.

After completing the vaccination task, participants completed the individual differences measures.

3. Results

To test H_1, we first investigated if there were any self-other discrepancies (indicated by an effect of target) and how they depended on the preferences of the other person (indicated by an interaction between target and information type). Next, to test H_2 to H_4, we investigated how the individual difference measures (H_2: numeracy, H_3: empathy, H_4: risk taking in health) moderated these effects. We conducted analyses separately for each of the dependent variables (worry, likelihood to vaccinate, and WTP). Analyses were conducted with SPSS 20. Because we did not set an upper limit on the WTP measure, some individuals indicated extremely large values that were influential outliers (e.g., mean WTP > 1000 Euros where the median WTP was 115). To correct for this, we winsorized the data by replacing values above the 95th percentile of the sample data with the value of the 95th percentile. [7]

---

[7] As an alternative approach, we trimmed the data by removing the extreme values (mean WTP>1000). Both approaches produced very similar results. Here we report the winsorized results using the full sample.

3.1. Self-other discrepancies

To investigate self-other discrepancies and how they varied depending on type of information (*H_1*), we used a repeated measures general linear model (GLM). For each dependent variable (worry, likelihood to vaccinate, and WTP), we tested a model with target and probability as repeated factors, and information type as between-subjects factors, controlling for the effect of order. To clarify interactions and test simple effects we used t-tests or post hoc comparisons, where applicable. All significant multivariate effects in the main analyses are reported ($p<.05$). Finally, we investigated whether self-other discrepancies were due to changes in participants' decisions for themselves or for the other. To do that, we followed up with repeated measures models separately for each target controlling for order.

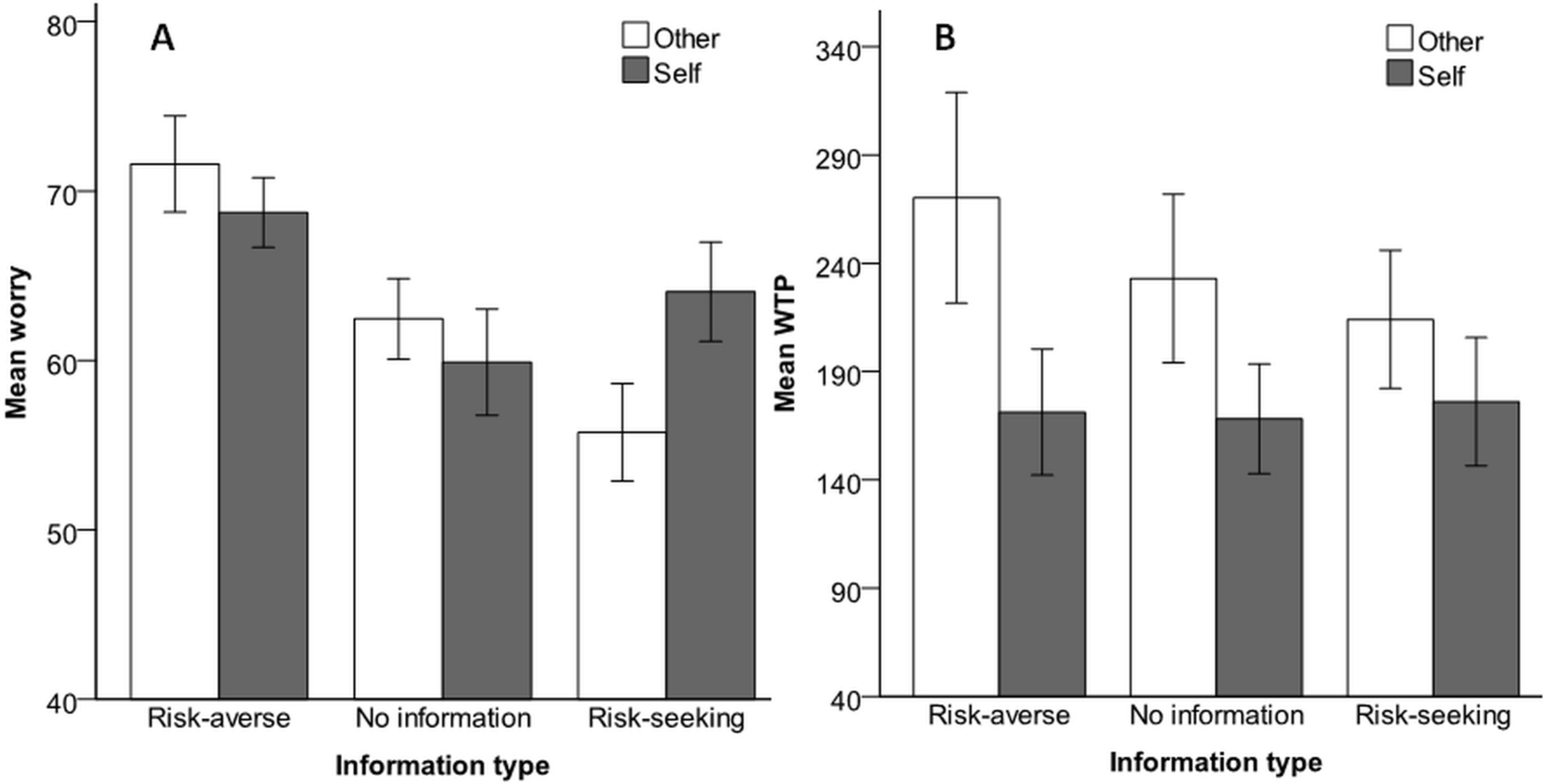
Self-other discrepancies in worry. Overall, the higher the probability of infection with the virus, the more worry participants reported, $F(6,135) = 103.88$, $p < .001$, $\eta_p^2 = .82$. An interaction between target and information type indicated that the direction of self-other discrepancies depended on the type of information participants received, $F(2, 140) = 5.34$, $p = .006$, $\eta_p^2 = .07$. Figure 1a shows that when participants received no information about the other person's preferences or received information that she was risk-averse, they predicted similar worry for the other person and themselves (no information: $M_{\text{other-self}} = 2.56$, $t(48) = .99$, $p = .325$; risk-averse: $M_{\text{other-self}} = 2.87$, $t(44) = 1.29$, $p = .205$). In contrast, when the other person was described as risk-seeking, they predicted that she would be less worried than they would be, $M_{\text{other-self}} = -8.31$, $t(49) = -2.43$, $p = .019$. Follow-up analyses for each target separately revealed that information type had no significant effect on how worried participants predicted themselves to be, $F(2, 140)=2.45$, $p=.090$, $\eta_p^2 =.03$. Instead, information type had an effect on how much worry participants predicted for the other person, $F(2, 140) = 8.98$, $p < .001$, $\eta_p^2 = .11$. When the other person was described as risk-averse, participants predicted that she would be more worried compared to when she was described as risk seeking (p < .001, see Figure 1a) or there was no information (p = .018). When the other person was described as risk seeking, participants predicted that she would be less worried than when there was no information but this difference was not significant (p=.073).In sum, participants showed self-other discrepancies in predicted worry in line with the information provided about the other person, and these discrepancies were produced by changes in the prediction for the other person as a function of the information provided.

Self-other discrepancies in likelihood to vaccinate. Overall, the higher the probability of infection with the virus, the more likely participants were to opt for vaccination, $F(6, 135) = 59.78$, $p < .001$, $\eta_p^2 = .73$. Moreover, participants were more likely to have their sister vaccinated ($M = 77$, $SD = 18$) than they were to vaccinate themselves ($M = 72$, $SD = 20$), $F(1, 140) = 12.56$, $p = .001$, $\eta_p^2 = .08$. However, this discrepancy did not vary significantly as a function of the information participants received, $F(2, 140) = .49$, $p = .612$, $\eta_p^2 = .01$. Means in all three conditions pointed towards a more risk-averse decisions for the sister: $M_{\text{other-self}} = 5.40$, $t(44) = 2.73$, $p = .009$for risk-averse vs. $M_{\text{other-self}} = 6.52$, $t(48) = 2.94$, $p = .005$ for no information, and $M_{\text{other-self}} = 3.18$, $t(49) = 1.08$, $p = .284$for risk seeking. In sum, participants were more likely to have their sister vaccinated than have themselves vaccinated, and, consistent with H_{1b}, this effect did not vary significantly as a function of the preferences of the sister.

Self-other discrepancies in WTP. Overall, the higher the probability of infection with the virus, the more participants were willing to pay for vaccination, $F(6, 135) = 16.41$, $p < .001$, $\eta_p^2 = .42$. There was a significant effect of target indicating self-other discrepancies in WTP, $F(1, 140) = 21.45$, $p < .001$, $\eta_p^2 = .13$. On average, participants were willing to pay around €66 more for the vaccination of the other person (M = €238, $SD = 275$) than for their own (M = €172, SD = 193) (see Figure 1b). This effect varied as a function of the probability of contracting the virus, $F(6, 135) = 2.54$, $p = .023$, $\eta_p^2 = .10$, such that the size of the self-other discrepancy was larger for larger probabilities (≥ 50%) $M_{\text{other-self}} = 89$, $SD = 232$, than for smaller probabilities (< 50%) $M_{\text{other-self}} = 37$, $SD = 112$, $t(143) = -3.90$, $p<.001$. The self-other discrepancy did not vary as a function of information type, $F(2, 140) = 1.45$, $p = .239$, $\eta_p^2 = .02$ (Figure 1b).

Means in all three conditions pointed towards more risk-averse decisions for the sister: $M_{\text{other-self}} = 99$, $t(44) = 2.99$, $p = .004$for risk-averse vs. $M_{\text{other-self}} = 65$, $t(48) = 2.58$, $p = .013$for no information, and $M_{\text{other-self}} = 38$, $t(49) = 2.46$, $p = .018$for risk seeking. In sum, consistent with H_{1b}, participants were willing to pay more to have their sister vaccinated than have themselves vaccinated, regardless of her preferences. This discrepancy was larger when the probability of contracting the virus was larger.

Figure 1. Mean predicted worry (Panel A) and mean willingness to pay (WTP, Panel B) as a function of target (self or other) and information type. Error bars are ±1 standard error of the mean.



3.2. Individual differences

To investigate how individual differences moderated self-other discrepancies (H_2 to H_4), we conducted similar GLMsas in the previous section by adding the three individual difference variables to the models (as continuous variables). In these analyses we controlled for the effect of gender of the participant. We thus conducted three analyses (on worry, likelihood, and WTP), with information type, numeracy, empathy, risk taking in health, order, and gender as independent variables. In particular, we investigated if the individual difference variables moderated the effects of target and the interactions between target and information type reported above. The model hence tested for two-way interactions between decision target and each of the individual difference measures, as well as for three-way interactions between target, information type, and each of the individual difference measures. In order to clarify significant interactions, we divided participants into high and low groups based on median splits of the individual difference variables and examined with follow-up t-tests how the effects differed between these groups.

In addition, we investigated whether self-other discrepancies were due to changes in participants' decisions for themselves or for the other. In particular, like in the previous section we ran analyses separately for each target and investigated if there were significant between-subjects effects of the individual difference variable, or a significant interaction between the individual difference variable and information type.

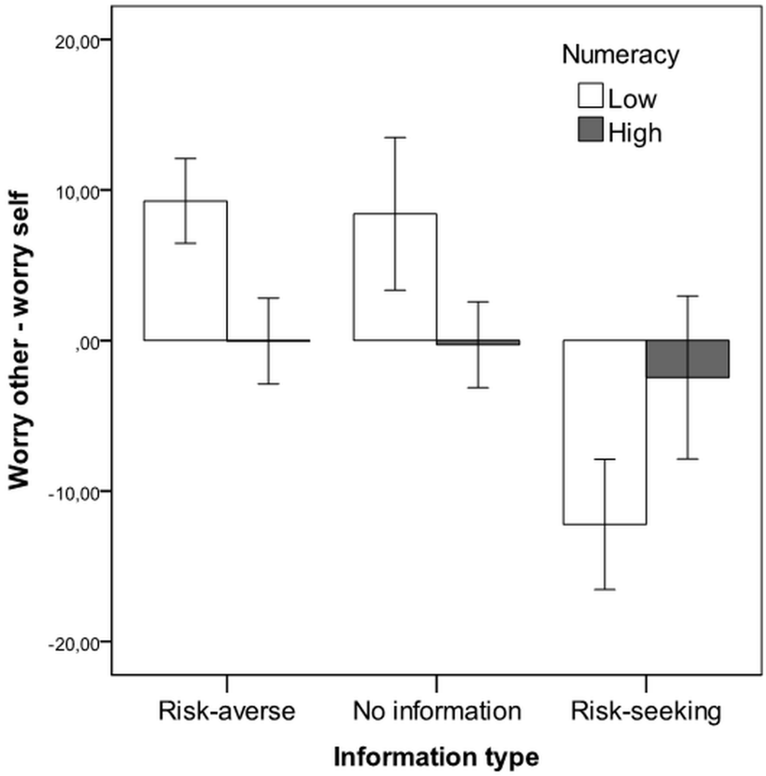
Individual differences in self-other discrepancies in worry. Numeracy moderated the effect of information type on self-other discrepancies, i.e., there was an interaction between numeracy, information type, and target, $F(2, 128) = 3.44$, $p = .035$, $\eta_p^2 = .05$. Figure 2 shows that participants with high numeracy did not show self-other discrepancies in predicted worry: they tended to predict similar worry for themselves and for the other person regardless of the information they received ($M_{\text{other-self}}$ = −.03, $t(30)$ = −.01, p =.992, for risk-averse; $M_{\text{other-self}}$ = −.28, $t(32)$ = −.10, p =.923, for no information; $M_{\text{other-self}}$ = −2.45, $t(19)$ = −.45, p =.656, for risk seeking). In contrast, participants with low numeracy showed large discrepancies, which were in line with the information they received about the other person ($M_{\text{other-self}}$ = 9.28, $t(13)$ = 3.30, p = .006,for risk-averse; $M_{\text{other-self}}$ = 8.41, $t(15)$ = 1.66, p = .118,for no information; $M_{\text{other-self}}$ = −12.21, $t(29)$ = −2.82, p = .009,for risk seeking). Follow-up analyses for each target separately revealed that numeracy had no significant effect on worry for the self, $F(1, 128)$ = 1.18, p = .280, η_p^2 = .01, and it did not interact with information type, $F(2, 128)$ = 1.04, p = .356, η_p^2 = .02. Numeracy had no significant effect on worry for the other, $F(1, 128)$ = 1.36, p = .246, η_p^2 = .01, and did not interact with information type, $F(2, 128)$ = .96, p = .388, η_p^2 = .02.

Risk taking propensity and empathy had no significant effects on self-other discrepancies in worry, $p > .1$.

In sum, consistent with H_2 applied to predictions of feelings, participants with low numeracy showed self-other discrepancies in predicted worry that reflected the other person's preferences, while participants with high numeracy predicted similar worry for themselves and the other person regardless of the information they received about her. However, numeracy had no direct effect on predicted worryfor the other suggesting that these self-other discrepancies were likely produced by relative adjustment of predictions for both targets.

Individual differences in self-other discrepancies in likelihood. There were no significant effects of any of the individual measures, $p > .1$.

Figure 2. Mean self-other discrepancy in predicted worry as a function of information type and numeracy. Mean discrepancy is worry other−worry self, where a score>0 indicates more predicted worry for the other person than for oneself. Low numeracy is indicated by a score<3 and high numeracy by a score ≥ 3 on the adaptive Berlin Numeracy Test. Error bars are ±1 standard error of the mean.



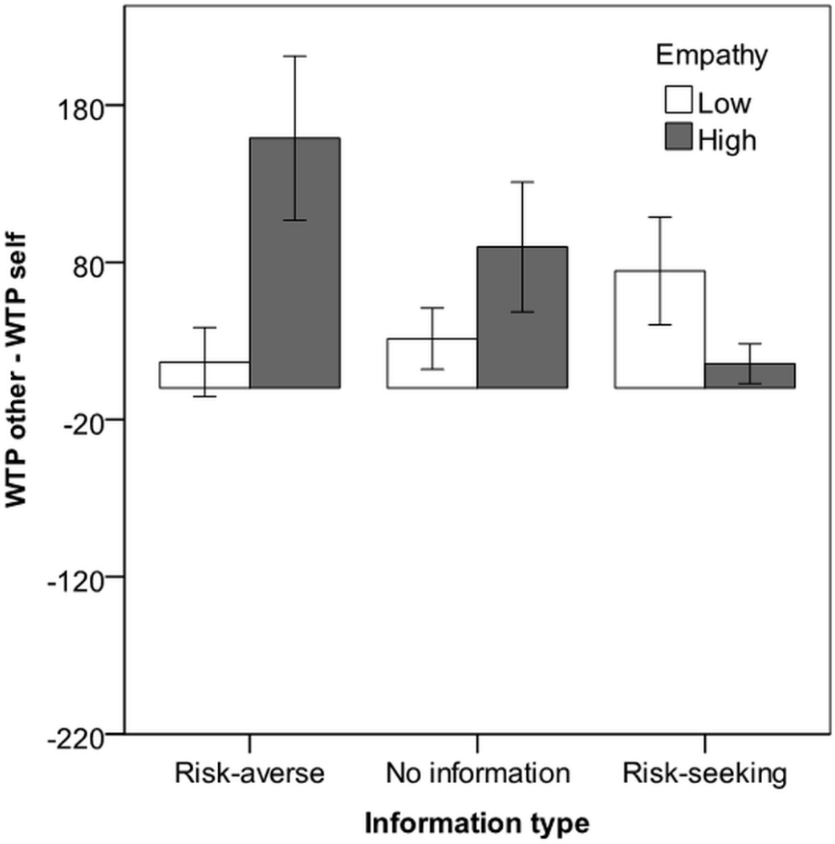
Individual differences in self-other discrepancies in WTP. Empathy moderated the effect of information type on self-other discrepancies, *F*(2, 128) = 3.72, *p* = .027, η_p^2= .06. Figure 3 shows that participants who were high in empathy tended to show self-other discrepancies more in line with the wishes of the other person. In contrast, relative to decisions for themselves, the WTP responses of participants low in empathy followed the other person's wishes to a lesser extent. To illustrate, when the other person was described as risk-averse, self-other discrepancies were larger for participants high vs. low in empathy ($M_{\text{other-self}}$ = 159 vs. $M_{\text{other-self}}$ = 17, *t*(33) = −2.52, *p* = .017). When the other person was described as risk-seeking, self-other discrepancies were smaller for participants high vs. low in empathy, although this difference was not significant ($M_{\text{other-self}}$ = 16 vs. $M_{\text{other-self}}$ = 75, $t(23)$ = 1.62, p = .120). Follow-up analyses separately for each target showed that empathy had no effect on WTP for the self, $F(1, 128)$ = 1.03, p = .312, η_p^2 = .01, and did not moderate the effect of information type on WTP for the self, $F(2, 128)$ = .30, p = .744, η_p^2 =.01. Further, empathy did not influence WTP for the other, $F(1, 128)$ = 2.53, p = .114, η_p^2 =.02, nor moderated the effect of information type, $F(2, 128)$ = 2.23, p =.111, η_p^2 =.03.

Numeracy showed no significant effects, $p > 1$.

Self-reported risk taking in the domain of health also moderated self-other discrepancies, $F(1, 128)$ = 5.38, p = .022, η_p^2 = .04. Participants who were risk takers (divided based on median split) showed larger self-other discrepancies, with a mean discrepancy score of $M_{\text{other-self}}$ =102 (SD = 209), than participants who tended to avoid risks, with a mean discrepancy score of $M_{\text{other-self}}$ = 30 (SD = 121), $t(114)$ = −2.56, p = .012. Follow-up analyses separately for each target showed that self-reported risk taking in the domain of health had no effect on WTP for the self, $F(1, 128)$ = 1.26, p = .264, η_p^2 = .01. However, self-reported risk taking in the domain of health had an effect on WTP for the other, $F(1, 128)$ = 5.07, p = .026, η_p^2 = .04, such that participants who reported being risk-takers themselves were especially likely to pay more for the vaccination of the other person (M = 282, SD = 308) compared to participants who were risk-averse (M = 194, SD = 232).

In sum, consistent with H_3, participants high in empathy were more likely to take into account the preferences of the other person relative to their own decisions, which resulted in a different extent of self-other discrepancies in WTP as a function of the surrogate’s empathy. However, empathy had no direct effect on decisions for the other, suggesting that these self-other discrepancies were likely produced by relative adjustment of decisions for both targets. In addition, partially consistent with H_4, participants who were risk takers themselves showed larger self-other discrepancies than participants who were not risk takers. However, these discrepancies did not stem from decisions for oneself but from decisions for the other person: risk takers were willing to pay more for vaccination of the other person compared to participants who avoided risks.

Figure 3. Mean self-other discrepancy in willingness to pay (WTP) as a function of information type and empathy. Mean discrepancy is WTP other−WTP self, where a score>0 indicates more willingness to pay for the other person than for oneself. Low vs. high empathy groups are based on median split. Error bars are ±1 standard error of the mean.



4. Discussion

Generally, our results supported *social values theory*. When participants had to make a health decision for themselves and for another person, they made more risk-averse decisions for the other person than for themselves. This was the case both when they had no information about what the other person might potentially prefer (no information condition), and when the preferences of the other person were in line with the presumed social norm (i.e., avoiding risks in the health domain) – a result that replicates findings from previous studies (Dore et al., 2014; Stone & Allgaier, 2008; Stone et al., 2013). Interestingly, when the preferences of the other person were counter-normative (i.e., risk-seeking), participants still showed self-other discrepancies in line with the social norm of risk-aversion: They were willing to pay more for the other person's vaccination than for their own. Our analysis of participants' feeling and their predicted feelings for the other person showed that these results could not be explained by self-other discrepancies in predictions (see also Stone et al., 2013). In particular, participants in the risk-seeking condition predicted that the other person would be less worried than themselves; however, they did not incorporate this prediction in their decisions.

These findings show that the perceived social value of risk predicts self-other discrepancies in risky decisions, even when the preferences of the other person are at odds with the social norm. This was the case regardless of what the other person preferred and even though the assumed responsibility of decision makers was fictitious. Research shows that regret concerns and desire to minimize blame lead to increased risk avoidance in decisions for other people (Atanasov, 2010; Stone, Yates, & Caruthers, 2002). Outside of the laboratory, where responsibility and consequences of decisions are real, the social or legal norms for decisions where risk is entailed can have an even stronger influence. For example, doctors tend to make more conservative decisions for their patients than they do for themselves and they often do so out of fear of legal prosecution (Garcia-Retamero & Galesic, 2012, 2014). Future research should investigate if this "norm-over-preferences" effect exists also in other domains, especially those where risk-seeking is valued (e.g., for example some social domains or casual romantic relationships, Beisswanger et al., 2003; Stone & Allgaier, 2008).

Previous research has established that factors like anxiety (Wray & Stone, 2005) or depression (Garcia-Retamero et al., 2015) moderate self-other discrepancies. To the best of our knowledge, our study is the first that shows how cognitive abilities, empathic concern for others, and risk taking propensity influence the extent to which people make different predictions or decisions for others and for themselves. In particular, this study examined the role of numeracy in self-other discrepancies. Numeracy is an important component of risk literacy:the ability to make good decisions based on numerical estimates of risk (Cokely et al., 2012; Galesic & Garcia-Retamero, 2010), and numerical risk information ismore common than ever in health decision making (Garcia-Retamero, Andrade, Sharit, & Ruiz, 2015). In our study, people with lower numeracy showed self-other discrepancies in emotions, while people with high numeracy predicted similar emotions for themselves and the other person. The self-other discrepancies in emotions in the predictions of participants with low numeracy were actually in line with the preferences of the other person, showing that low numeracy individuals were more likely to incorporate these into their predictions. High numeracy individuals, on the other hand, tend to base their feeling on the exact risk estimates or number comparisons (Peters, 2012; Petrova et al., 2014) rather than on more general factors of the situation (Garcia-Retamero & Galesic, 2010; Peters et al., 2006; Västfjäll et al., 2011). This might have eliminated self-other discrepancies in the predictions of high numeracy individuals, as they may have been more focused on the numerical risks, which were analogous in predictions for themselves and predictions for the other person. We should also note that this difference in predicted emotions did not translate into different decisions, suggesting that low and high numeracy individuals used the social norm and preferences to a similar extent.

Although self-other discrepancies were generally consistent with the social value of risk, the extent of the discrepancy varied as a function of participants' empathy. Participants who reported high tendencies for empathic concern for others tended to make decisions for others that were more consistent with what others wanted, relative to decisions for themselves. To illustrate, when the other person had risk-averse preferences, individuals high on empathy were willing to pay a lot more for the vaccination of the other person than for their own, thus in a way incorporating the other's preferences. When the other person was risk-seeking, individuals high on empathy showed smaller self-other discrepancies compared to individuals low on empathy; those low on empathy on average seemed to follow the social norm and would pay more for the other person's vaccination (see Figure 3). In other words, relative to their own decisions, participants high on empathy were more likely to incorporate the target's preferences than participants low on empathy, even when these preferences were at odds with the social norm. Previous research has demonstrated that when empathy is facilitated people predict that others will make decisions similar to their own (Faro & Rottenstreich, 2006; Garcia- Retamero et al., 2015; Hsee & Weber, 1997). Our research, in turn, shows that empathy can also lead to larger self-other discrepancies in decisions, and that these discrepancies may stem from trying to fulfill the wishes of the other person. Overall, people high in empathy may be more likely to follow both social norms and the person's wishes when they make decisions for others relative to decisions for themselves.

Participants who were themselves risk takers in the domain of health showed larger self-other discrepancies than participants who were not risk takers. However, surprisingly, participants' self-reported risk taking was not significantly related to their willingness to pay for their own vaccination. It is possible that vaccination is a behavior that shows little variance between individuals, such that the vast majority of individuals vaccinate, as opposed to other risky behaviors captured by the DOSPERT scale in which individuals may be more likely to vary (e.g., drinking at a social function, not applying sunscreen). This could explain why self-reported risk takers were not willing to pay less for their own vaccination compared to those who were less likely to engage in health and safety risks in general. Instead, risk taking had an effect on decisions for the other. Consistent with social values theory, when people perceive that their decisions for themselves coincide with the social norm (i.e., they report to be risk averse), they make similar decision for other people, thereby diminishing self-other discrepancies. For instance, the large discrepancies in decisions of participants who were risk takers could result from counter-projection (Fagerlin et al., 2001). The decisions these participants made for themselves were similar to the decisions of people who were risk averse. However, risk takers may be aware of their frequently counter-normative risky decisions, and hence might have been motivated to make a decision for the other person that they perceived as consistent with social norms, thereby resulting in even more risk-averse decisions for others, resulting in larger self-other discrepancies

One limitation of the current research is that there were no effects of numeracy or empathy on decisions for others, as we predicted. Follow up analyses on the self-other discrepancies that were a function of numeracy and empathy showed no significant effects neither on decisions for the self, nor on decisions for the other. It is possible that these discrepancies were produced by some relative adjustment of judgments (i.e., anchoring), for example depending on the order in which participants decided for themselves and for the other person. However, we did not discover any informative patterns or significant effects, possibly due to low power to detect such differences. Alternatively, these self-other discrepancies may be due to Type I error. Future research should try to replicate our findings and investigate when self-other discrepancies in decisions result from changes in decisions for the other or changes in decisions for oneself.

Finally, we should note that we did not directly assess the perceived social value of risk in the domain of health. However, previous research in similar populations using similar scenarios (e.g., vaccination and flu outbreaks) has established that in domains where health and physical safety are involved, risk-aversion is the more socially acceptable option (Dore et al., 2014; Stone et al., 2013), providing support to our assumption. Nevertheless, the possibility remains that not the perceived social norm per se but more specifically people's desire to avoid possible poor outcomes associated with the decision drive the observed effects. For example, one could speculate that a diminishing responsibility for the decision target could be related to a smaller influence of the social norm (because of a smaller likelihood of punishment or blame in the case of a poor outcome) and thus potentially more influence of the preferences of the target. It is also possible that social norms evolve at least partially as a result of people's desire to avoid such poor outcomes, resulting in a possible overlap between the two notions. It remains for future research to disentangle the exact motivations behind people's decisions for others and the resulting discrepancies with decisions they make for themselves in the domain of health.

Given the number of situation− and person−based factors that play a role in self-other discrepancies, it is no wonder that the literature has shown some mixed results (see Garcia-Retamero et al., 2015). However, recent theoretical and empirical advances have started shedding light on the processes behind self-other discrepancies (Garcia-Retamero et al., 2015; Stone et al., 2013). Ultimately, one of the goals of this line of research is to inform surrogate decision making and increase accuracy. However, what is considered an accurate surrogate decision can vary depending on the standard that is adopted: advance directive, substituted judgment or best interest (Lawrence & Brauner, 2009). For example, the advance directive requires that surrogates follow the preferences stated by the decision target, while the substituted judgment requires surrogates to make a decision that the decision target would have made if able. Thus, accurately predicting the target's feelings towards risks and options and the ability to incorporate these into decisions are potentially essential in surrogate decision making. Our results show that individual characteristics of surrogates like numeracy, empathic concern or propensity towards risk taking, and the willingness to follow contra-normative preferences can potentially influence the accuracy of surrogate decisions. For instance, intuitively, a more empathic family member may be a better surrogate decision maker than a less empathic one if substituted judgment is to be followed. Research in a more ecological setting can investigate to what extent relevant individual differences influence surrogate accuracy and what implications these have for stakeholders.

This research showed how a host of factors (informational, cognitive, and emotional) influence to what extent our own decisions about risk are different from the decisions we make for other people, and to what extent these differences are based on social norms. Overall, results confirmed that self-other discrepancies at least partially result from following social norms when deciding for others. In this experiment, this effect persisted even when the wishes of the other person were known to be different from what social norms dictated and was larger when decision makers were less empathic.

5. References

Atanasov, P. D. (2010). Double risk aversion. *Available at SSRN. http://papers.ssrn.com/sol3/papers.cfm?abstract_id = 1682569* (Last accessed 23 January, 2015).

Beisswanger, A. H., Stone, E. R., Hupp, J. M., & Allgaier, L. (2003). Risk taking in relationships: Differences in deciding for oneself versus for a friend. *Basic and Applied Social Psychology, 25*(2), 121-135.

Blais, A., & Weber, E. U. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgment and Decision Making, 1*(1), 33-47.

Brody, J. L., Annett, R. D., Scherer, D. G., Perryman, M. L., & Cofrin, K. M. (2005). Comparisons of adolescent and parent willingness to participate in minimal and above-minimal risk pediatric asthma research protocols. *Journal of Adolescent Health, 37*(3), 229-235.

Butler, S., Rosman, A., Seleski, S., Garcia, M., Lee, S., Barnes, J., & Schwartz, A. (2012). A medical risk attitude subscale for DOSPERT. *Judgment and Decision Making*, *7*, 189–195.

Chapman, G. B., & Coups, E. J. (2006). Emotions and preventive health behavior: worry, regret, and influenza vaccination. *Health Psychology*, *25*, 82-90.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making, 7*, 25-47.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.

Dieckmann, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis, 29*(10), 1473-1488.

Dore, R. A., Stone, E. R., & Buchanan, C. M. (2014). A social values analysis of parental decision making. *The Journal of Psychology: Interdisciplinary & Applied*, *148*(4),477-504.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*(3), 327.

Fagerlin, A., Ditto, P. H., Danks, J. H., & Houts, R. M. (2001). Projection in surrogate decisions about life-sustaining medical treatments. *Health Psychology, 20*(3), 166-175.

Faro, D., & Rottenstreich, Y. (2006). Affect, empathy, and regressive mispredictions of others' preferences under risk. *Management Science, 52*(4), 529-541.

Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine*, *170*, 462-468.

Galesic, M., & Garcia-Retamero, R. (2011). Do low-numeracy people avoid shared decision making? *Health Psychology, 30*(3), 336-341.

Garcia-Retamero, R., Andrade, A., Sharit, J., & Ruiz, J. G. (2015). Is patient's numeracy related to physical and mental health? *Medical Decision Making*, *35*, 501-511.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science*, *22*, 392-399.

Garcia-Retamero, R., & Galesic, M. (2010). How to reduce the effect of framing on messages about health. *Journal of General Internal Medicine, 25*(12), 1323-1329.

Garcia-Retamero, R., & Galesic, M. (2011). Using plausible group sizes to communicate information about medical risks. *Patient Education and Counseling, 84*(2), 245-250.

Garcia-Retamero, R., & Galesic, M. (2012). Doc, what would you do if you were me? on self–other discrepancies in medical decision making. *Journal of Experimental Psychology: Applied, 18*(1), 38-51.

Garcia-Retamero, R., & Galesic, M. (2014). On defensive decision making: How doctors make decisions for their patients. *Health Expectations, 17*(5), 664-669.

Garcia-Retamero, R., Okan, Y., & Maldonado, A. (2015). The impact of depression on Self–Other discrepancies in decision making. *Journal of Behavioral Decision Making, 28*(1), 89-100.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Hsee, C. K., & Weber, E. U. (1997). A fundamental prediction error: Self–others discrepancies in risk preference. *Journal of Experimental Psychology: General, 126*(1), 45-53.

Kray, L. J. (2000). Contingent weighting in self-other decision making. *Organizational Behavior and Human Decision Processes, 83*(1), 82-106.

Kray, L., & Gonzalez, R. (1999). Differential weighting in choice versus advice: I'll do this, you do that. *Journal of Behavioral Decision Making, 12*(3), 207-218.

Lawrence, R. E., & Brauner, D. J. (2009). Deciding for others: limitations of advance directives, substituted judgment, and best interest. *Virtual Mentor*, *11*(8), 571.

Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. *Health Psychology, 24*(4S), S49-S56.

Marks, M. A., & Arkes, H. R. (2008). Patient and surrogate disagreement in end-of-life decisions: Can surrogates accurately predict patients' preferences? *Medical Decision Making : An International Journal of the Society for Medical Decision Making, 28*(4), 524-531.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31-35.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*(5), 407-413.

Petrova, D., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making, 27*, 191-199.

Shalowitz, D. I., Garrett-Mayer, E., & Wendler, D. (2006). The accuracy of surrogate decision makers: A systematic review. *Archives of Internal Medicine, 166*(5), 493-497.

Stone, E. R., & Allgaier, L. (2008). A social values analysis of self–other differences in decision making involving risk. *Basic and Applied Social Psychology, 30*(2), 114-129.

Stone, E. R., Choi, Y., de Bruin, W. B., & Mandel, D. R. (2013). I can take the risk, but you should be safe: Self-other differences in situations involving physical safety. *Judgment and Decision Making, 8*(3), 250-267.

Stone, E. R., Yates, A. J., & Caruthers, A. S. (2002). Risk taking in decision making for others versus the self. *Journal of Applied Social Psychology, 32*(9), 1797-1824.

Västfjäll, D., Peters, E., & Starmer, C. (2011). *Numeracy, Incidental Affect, and the Construction of Prices, (Report No. 11-01).* Eugene: Decision Research.

Weinstein, N. D. (1984). Why it won't happen to me: Perceptions of risk factors and susceptibility. *Health Psychology, 3*(5), 431-457.

Wray, L. D., & Stone, E. R. (2005). The role of self-esteem and anxiety in decision making for self versus others in relationships. *Journal of Behavioral Decision Making, 18*(2), 125-144.

SECTION II:

FIELD RESEARCH

CHAPTER 6:

Numeracy predicts risk of pre-hospital decision delay

The content of this chapter has been submitted as Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E., Heredia Carrasco, A., Arrebola Moreno, A., & Ramírez Hernández, J. A. Numeracy predicts risk of pre-hospital decision delay. A retrospective study of acute coronary syndrome survival.

Chapter 6

Numeracy Predicts Risk of Pre-Hospital Decision Delay: A Retrospective Study of Acute Coronary Syndrome Survival

Many patients delay seeking medical attention for acute coronary syndrome (ACS), profoundly increasing their risk for death and major disability. Although research has identified several risk factors, efforts to improve patient decision making have generally been unsuccessful, prompting a call for more research into psychological factors. We aimed to estimate the relations between ACS decision delay, known risk factors, and psychological factors with emphasis on general decision making skills that are often associated with health outcomes (i.e., numeracy—a strong predictor of one's ability to evaluate and understand risk). About five days after experiencing ACS, 102 survivors (mean age=58, 32-74) agreed to share health and medical information, and completed a questionnaire including measures of numeracy, decision delay, and other relevant factors (e.g., anxiety and depression, symptom severity, knowledge, demographics). Patient numeracy uniquely predicted decision delay, OR=.62 [95% CI .40, .95]. Independent of the influence of all other assessed factors, a patient with high (vs. low) numeracy was about four times more likely to seek medical attention within the critical first hour after symptom onset (i.e., 3.84 [1.127, 11.65]). Low numeracy appears to be one of the largest decision delay risks identified to date. Even small differences in numeracy may confer protective benefits that far exceed those of other education-related factors (e.g., knowledge of ACS symptoms). Results accord with theories emphasizing influences of patient deliberation, denial, intentions, and outcome-understanding during delay decision making. Findings suggest new ways to identify and educate patients at higher risk for life-threatening decision delay using brief numeracy tests and custom-tailored risk communications.

1. Introduction.

Cardiovascular disease is the number one killer worldwide responsible for about one in three deaths (World Health Organization, 2014). Almost half of all deaths from cardiovascular disease result from acute coronary syndromes (ACSs) (Turpie, 2006; World Health Organization, 2014). ACS is an umbrella term for a group of conditions characterized by an abrupt reduction in coronary blood flow that usually presents with chest pain, pain in one or both arms, and shortness of breath among other symptoms (Amsterdam et al., 2014). ACS survival rates and outcomes are relatively good when treatment is administered within one hour of initial symptom onset (Berger et al., 1999; Goldberg et al., 1992; Goldberg et al., 1998; Newby et al., 1996; Perkins-Porras et al., 2009). Unfortunately, up to 80% of patients who receive treatment do not receive it within the critical first hour (DeVon, Hogan, Ochs, & Shapiro, 2010). Estimated median pre-hospital treatment delays range from at least two to four hours, with more than a third of these patients failing to receive treatment within the first 6 hours when some myocardial salvage is still possible (DeVon et al., 2010). Why do so many people risk long delays during ACS?

Although there are many factors that affect time to treatment, an estimated two-thirds of the delay time results because patients decide to wait instead of immediately seeking medical help (e.g., *decision delay* (Mackay et al., 2014)). This decision to delay treatment is among the most life-altering decisions that people make during ACS, profoundly increasing their risk for serious complications, major disability, and death (Goldberg et al., 1998; Mackay et al., 2014; Moser et al., 2007). To date, most decision delay research has investigated demographic, situational, and disease-related factors. Those who believe they are experiencing a heart attack and those who report having severe symptoms are about three times more likely to receive treatment within an hour. Those who are married, who experience ACS outside of the home, and who are younger tend to be nearly twice as likely to receive treatment within an hour (e.g., estimates range from around 60%-220% more likely, Nguyen et al., 2010; Perkins-Porras et al., 2009; Smolderen et al., 2010).

Educational efforts generally aim at empowering patient decision making during ACS by focusing on symptom knowledge and the importance of quick action. Despite the use of extensive, well-designed, and long-term educational and media campaigns nearly every published intervention has failed to reduce decision delay (Mooney et al., 2012). Even randomized control trials utilizing individual patient counseling and education techniques have proven ineffective (for some recent exceptions see Mooney et al., 2012; 2014). In response to these failures and other findings experts suggest that reducing the burden of the disease requires more research on the psychological factors that shape ACS decision making. In particular, an understanding of decision processes and skills may be useful for individualized interventions and identification of those who are at the greatest risk (Moser et al., 2007).

Numeracy and Health Outcomes. One skill that is potentially relevant to decision delays is a person's practical ability to solve problems involving probability (i.e., statistical numeracy), which tends to be a robust predictor of diverse health and medical outcomes (Garcia-Retamero & Galesic, 2013; Peters, Hibbard, Slovic, & Dieckmann, 2007; Peters, 2012; Reyna, Nelson, Han, & Dieckmann, 2009). Patients with lower numeracy tend to have a higher prevalence of comorbidities, take more prescribed medications, and have more negative perceptions of their health (Garcia-Retamero, Andrade, Sharit, & Ruiz, 2015). Patients with low numeracy are also more likely to be hospitalized and visit emergency services (Apter et al., 2006; Ginde, Clark, Goldstein, & Camargo, 2008). Among patients with diabetes, low numeracy is associated with less self-efficacy, fewer self-management behaviors, and poorer glycaemic control (Cavanaugh et al., 2008; Marden et al., 2012; Osborn, Cavanaugh, Wallston, & Rothman, 2010). Chronic kidney disease patients with low numeracy are less likely to receive a kidney transplant or be on the waiting list for one (Abdel-Kader et al., 2010). Lower numeracy has also been associated with difficulties in proper use of contraception among urban women (Yee & Simon, 2014) and poorer HIV medication management among patient on antiretroviral therapy (Waldrop-Valverde, Jones, Gould, Kumar, & Ownby, 2010).

Emerging research suggests that numeracy is also a predictor of cardiovascular decision making and health outcomes. Small but significant increases in prevalence of myocardial infarction, congestive heart failure, and diabetes have been found among male patients at a veteran's clinic in the U.S. (Garcia-Retamero et al., 2015). Among patients with acute heart failure who presented to an emergency department, those who had low self-reported numeracy were 40% more likely to return to the hospital in the following months (McNaughton et al., 2013). Among patients taking warfarin, low numeracy was associated with poor anticoagulation control (Estrada, Martin-Hryniewicz, Peek, Collins, & Byrd, 2004). Additionally, patients with lower numeracy were less likely to correctly remember the risks of heart disease associated with being overweight and exercising (Galesic & Garcia-Retamero, 2011). More generally patients' memory for risk and benefit information related to treatment options for angina like balloon angioplasty is also linked to numeracy (Zikmund-Fisher, Smith, Ubel, & Fagerlin, 2007). Numeracy also predicts whether or not patients will consider meaningful value comparisons between different health states (Zikmund-Fisher et al., 2007).

A primary reason why numeracy tends to be a robust predictor of health outcomes is because numeracy is strongly linked to general decision making skills (e.g., fewer biases and errors), including the ability to understand and evaluate risk—i.e., risk literacy (see riskliteracy.org for examples; (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Peters, 2012; Reyna et al., 2009); see also health literacy, Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011; Sørensen et al., 2012). Numerate people also tend to become better informed decision makers because they are more likely to realistically evaluate and deliberate about decision benefits, risks, and trade-offs, more precisely integrating emotion and cognition into an intuitive understanding of decision options (Cokely et al., 2012; Cokely & Kelley, 2009; Garcia-Retamero, Cokely, Wicki, & Joeris, 2016; Peters et al., 2006; Peters, 2012; Reyna, 2012; Reyna et al., 2009). Theoretically, during ACS the decision to seek medical attention may involve many numeracy-related skills such as estimating severity/intensity and identifying/evaluating sources of risk efficiently and in a short amount of time. Indeed, appraisal of symptoms as urgent or serious, and a correct attribution of symptoms to a heart problem are two factors that are most strongly associated with shorter decision delay times (Moser et al., 2007). The decision to seek medical attention for ACS also involves weighing of potential benefits and costs (e.g., large benefits of life-saving immediate treatment for ACS compared to modest costs of suffering embarrassment and troubling others in case of a false alarm). In sum, we hypothesize that numeracy predicts better decision making in the event of ACS, such that patients with higher numeracy decide to seek medical attention faster. To test this hypothesis and model the interplay of relations between numeracy and other influential risk factors, we conducted a retrospective questionnaire study of post-acute coronary syndrome patients' pre-hospital decision delay.

2. Methods

2.1. Design

Participants were consecutive patients admitted to the Cardiology Department of the University hospital Virgen de las Nieves who were eligible and agreed to participate. The Ethics Committee of the Hospital Virgen de las Nieves of the University of Granada approved all data collection, which took place between January 2014 and July 2015 and followed similar recruitments procedures to previous studies investigating the role of psychological factors in this population (Arrebola Moreno et al., 2014). Inclusion criteria were 1) having elevated cardiac markers (i.e., Troponin-I) and having ischemia, and 2) being fluent in Spanish. The exclusion criteria were having an inflammatory disease or having neurological problems. We planned to recruit 100 patients based on power analysis with G*power (gpower.hhu.de) for multiple linear regression with alpha=.05, power=.85, and a moderate effect size=.2.

Participants who met inclusion criteria for the study completed a questionnaire containing demographic questions, questions about the cardiac event, and measures of symptom knowledge, numeracy, and hospital scale of anxiety and depression. Participants completed the questionnaire on average 5 days (±5 SD) after the cardiac event. A researcher informed clinically stable, eligible patients about the study in oral and written form, and informed consent was obtained. Special care was taken to minimize participant exclusion due to fatigue, illiteracy, or other reading difficulties. The researcher (DP or AHR) offered assistance with the questionnaire to all patients and gave detailed instructions when needed. Patients who declined participation due to feeling sick at the moment but agreed to be re-contacted were re-invited during the following days.

2.2. Survey and Assessment Measures

Demographics. Participants indicated their age, gender, education level (1=less than primary school, 2=primary school, 3=secondary school, 4=professional education, 5=university education), marital status (1=married, 2=living with a partner, 3=single, 4=divorced, 5=separated, 6=widowed), and residence (village or city).

Decision delay. Time from symptom onset to the decision to seek medical attention (e.g., call an ambulance or go to the hospital). Participants were asked to indicate at what time their symptoms started (t1) and at what time they decided to seek medical attention (t2). The number of minutes elapsed between t1 and t2 was calculated.

Travel time. Time from the decision to seek medical attention (t2) to the arrival at the hospital (t3). The number of minutes elapsed between t2 and t3 was calculated. Decision and travel times were verified by family members.

Number of symptoms. Patients were asked to list the symptoms they were experiencing. The total number of symptoms was calculated.

Modified Responses to Symptoms Questionnaire (MRSQ) (Burnett, Blumenthal, Mark, Leimberger, & Califf, 1995). A retrospective multiple choice questionnaire containing questions about the circumstances around the onset of symptoms. Patients indicated 1) where they were when symptoms started, 2) with whom they were, 3) what was the first thing they did when they noticed the symptoms, 4) what was the first thing others did when the patient told them about the symptoms, 5) what they thought the problem was (i.e., symptom attribution), 6) in what vehicle they reached the hospital, and 7) how severe they thought the symptoms were at onset (from 1=not at all severe to 6=very severe).

Symptom knowledge. The ACS response index was administered (Riegel et al., 2007). It lists 21 symptoms and for each symptom the patients has to indicate whether they thought it is a symptom of a heart attack (yes/no) or they did not know. The final score is a sum of the number of correct symptom classifications.

Numeracy. We measured both objective and subjective numeracy, as recent works suggests that although correlated, they may assess different underlying numeric competencies (e.g., objective measures more strongly linked to one’s ability to perform mathematic operations related to risk evaluation; subjective measures offer some unique assessment of numerical processing motivation and perceived self-efficacy with numbers, (Peters & Bjalkebring, 2014)).

Objective numeracy was measured with a statistical numeracy math test that has been used extensively in medical decision making research (e.g., "In a lottery, the chance of winning a car is 1 in 1,000. What percent of tickets in the lottery win a car?", Schwartz, Woloshin, Black, & Welch, 1997). We chose this test because it was short and relatively easy, and thus suitable for this patient population where low education is prevalent (see also Cokely et al. (2012; Cokely, Ghazal, & Garcia-Retamero, 2014) for a review of existing numeracy tests and their suitability in various contexts). Subjective numeracy was measured using self-report responses (Fagerlin et al., 2007). The subjective numeracy scale used was one initially developed as a proxy for objective numeracy, showing moderate correlations with objective measures ($r<.5$) with similar predictive power for some tasks (Fagerlin et al., 2007; Zikmund-Fisher et al., 2007). The subjective numeracy scale tends to have the advantage that it is also perceived as less burdensome by patients and may also reduce math test anxiety (Fagerlin et al., 2007). Specifically, this scale measures people's self-reported numerical abilities and preferences for numerical information on a 6-point Likert scale (e.g., How good are you at figuring out how much a shirt will cost if it is 25% off? (Not at all good – Extremely good)). One question from the original scale about a 15% tip was not used due to low cultural applicability (see Garcia-Retamero & Galesic, 2013). The items of both the objective and subjective measures were adapted to the Spanish context.

Doctor visits. Patients indicated the number of times they went to the doctor in the past year.

Mental state. Patients filled the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983). The scale consists of 14 questions that measure the levels of anxiety and depression the patient is experiencing (e.g., Worrying thoughts go through my mind, 3=a great deal of the time, 2=a lot of the time, 1=from time to time but not too often, 0=only occasionally). The final score was the sum of all questions, where a larger score indicates greater anxiety and depression.

2.3. Medical Record Information and Metrics

All the following measures were obtained from patients' medical records.

Cardiovascular disease history. We recorded whether patients had previous cardiovascular disease history (e.g., previous myocardial infarction or ischemic disease).

Smoking. Patients were classified as non-smokers, smokers, or ex-smokers by the hospital personnel.

Hypertension. Patients were classified as hypertensive if they (a) had an average systolic blood pressure of>140 mm Hg and/or a diastolic blood pressure of >90 mm Hg at rest, (b) had previous history of hypertension, or (c) were taking antihypertensive drugs.

Diabetes. Patients were classified as diabetics if they (a) had fasting blood glucose of >126 mg/dL, (b) had previous history of diabetes mellitus, or (c) were taking insulin/oral hypoglycemics.

Body-mass index (BMI). We computed patients' body-mass index as weight (kg)/ height (m) squared.

STEMI. We recorded whether participants suffered a ST-segment elevation myocardial infarction (STEMI) or a non-STEMI. STEMI reflects an acute coronary occlusion and indicates a poor early-term clinical outcome. In contrast, a non-STEMI usually reflects a critically ill but not complete coronary artery obstruction and indicates a poor long-term clinical outcome (Park et al., 2013).

Number of obstructed arteries. A coronary stenosis of more than 50% in a main branch (i.e., left main, left anterior descending, circumflex, or right coronary artery) was considered an obstructed vessel.

Number of stents. Number of coronary stents that the patient had implanted during a percutaneous coronary intervention. In case of previous ACSs, we only considered stents that were implanted as a consequence of the current ACS.

3. Results

A total of 188 eligible patients were approached. Of these, 11% (20) declined to participate in the study. Of the remaining 168 who agreed to participate, 69% (116) returned a completed questionnaire and had a full medical record. One patient was excluded due to death during hospitalization and 8 were excluded because the final diagnosis at discharge was determined not to be ACS. Five more patients were excluded because they were unable to provide decision times and other details of the cardiac episode (e.g., the patient had a cardiac arrest followed by memory loss). Thus, the final sample size was 102.

3.1. Patient characteristics

Patients were on average 58 years old (±9 SD) and 84% were male. Seventy-two percent were married or lived with a partner, 10% were single, 13% were divorced or separated, and 6% were widowed. Twenty percent had no education, 40% had only finished primary school, 11% had finished secondary school, 18% had received professional education, and only 10% had completed university. Forty-five percent lived in the city of Granada (Spain) and the remaining 55% in the nearby villages.

Thirty-two percent had a previous history of CV disease. Twenty-eight percent were non-smokers, 48% were smokers, and 24% were ex-smokers. Forty-six percent were hypertensive and 31% had diabetes. BMI ranged from 18 to 45, with a mean of 28 (±4) falling into the overweight category. Forty-five percent of patients had a STEMI. The mean number of obstructed arteries was 1.6 (±.97) and the mean number of stent implants was 1.3 (±1.08).

3.2. Cardiac event

Patients reported on average 3 symptoms (±1), the most prevalent being chest pains. The majority of patients were at home when symptoms started (66%) and were with their partner (47%) or alone (33%). On noticing the symptoms, 37% adopted an avoidance strategy (e.g., prayed that symptoms will go away, tried not to pay attention to them), 28% told someone, 15% self-medicated (took a pill or herbal remedy), and 16% sought medical attention in some way. Forty-nine percent of patients recognized that their problem was heart-related. The majority of patients arrived at the hospital in ambulance (46%) or by car (40%). The median travel time (t3-t2) to the hospital was 30 minutes.

Participants described their symptoms as moderately severe with a mean of 3.7 (±1.7). ACS symptom knowledge was low on average, with patients correctly classifying about 10 (±4) symptoms of 21. Patients reported going to the doctors on average 6 (±7) times in the past year. The average HADS score was 12 (±7), ranging from 1 to 35, showing that on average the sample showed low anxiety and depression scores.

3.3. Numeracy

On the objective numeracy test most of the patients answered 0 (50%) or 1 question correctly (27%), and only 24% answered correctly 2 (16%) or 3 (8%) questions. Mean objective numeracy was .8 (±1) and mean subjective numeracy was 22 (±9). There was a moderate significant correlation between objective and subjective numeracy, r=.27, p=.006.

Higher education was related to both objective (r=.44, p<.001) and subjective numeracy (r=.49, p<.001). Male patients had both higher objective, .90 (±1) vs .38 (±.62), t(100)=1.99, p=.049, and subjective numeracy scores, 23 (±8) vs 15 (±9), t(98)=4.07, p<.001. Fewer obstructed arteries were associated with higher objective (*r*= -.21, *p*=.036) and subjective numeracy (*r*= -.20, *p*=.051). There were no further significant associations of numeracy with other measures, classical CV risk factors, or items from the MRSQ. Symptom knowledge, the number of days that had elapsed from the CV event to the questionnaire, the travel delay, and the perceived severity of symptoms were unrelated to either of the numeracy measures (p>.05).

3.4. Associations between numeracy and decision delay

Decision delay times ranged from 0 minutes (i.e., patient reported seeking medical attention immediately after noticing the symptoms) to 5 months (i.e., patient had persistent symptoms that eventually culminated in ACS), with a median of 75 minutes. Due to a strongly skewed distribution, the raw times in minutes were split into quintiles.[8] Table 1 shows the decision delay quintiles. To express the relationship between numeracy and decision delay, we computed odds ratios for one unit change in the numeracy scale (OR). To provide a more intuitive measure of effect size, we also computed odds ratios for a change across the whole range of the scale, i.e., comparing highest vs. lowest numeracy on the reverse decision delay score (OR_{range}).

Simple ordinal regression with decision delay quintiles as dependent variable and objective numeracy as predictor showed that objective numeracy was significantly related to decision delay. Patients with higher objective numeracy showed shorter decision delays than those with lower objective numeracy, OR=.64 [95% CI .44, .92], *Chi-square*(1)=5.67, *p*=.017. Compared to a patient with lower levels of objective numeracy, a patient with high objective numeracy was four times more likely, OR_{range}=3.84 [1.127,

---

[8] As an alternative strategy, we log-transformed the raw decision times, which gave very similar results. We chose to report the results using quintiles in order to facilitate comparisons with previous literature, where mostly categorical measures of decision delay have been used (e.g., Nguyen, Saczynski, Gore, & Goldberg, 2010; Perkins-Porras, Whitehead, Strike, & Steptoe, 2009; Smolderen et al., 2010).

11.65], to seek medical attention within the first 50 minutes of symptom onset.

A similar regression with subjective numeracy as a predictor revealed a near-significant marginal trend wherein higher subjective numeracy predicted shorter decision delay, OR=.96 [95% CI .93, 1.004], *Chi-square*(1)=3.10, *p*=.076. Compared to a patient with lower subjective numeracy, a patient with high subjective numeracy was about three times more likely to seek medical attention within 50 minutes of symptom onset, OR_{range}=2.97 [.89, 10.06].

Table 1. Decision delay quintiles (in minutes).

Decision delay	N	%
<= 20	24	23.5
21 - 50	17	16.7
51 - 180	21	20.6
181 - 750	20	19.6
751+	20	19.6
Total	102	100.0

Multiple regression analysis. We tested a multiple regression model with decision delay quintiles as dependent variable and objective and subjective numeracy as predictors. In this analysis we also entered several control variables (see Table 2). The model was significant, *Chi-square*(10)=24.70, *p*=.006, indicating that objective numeracy was more consistently related to decision delay than subjective numeracy, which failed to provide any unique explanatory power. Adding the control variables exerted a trivial, non-significant influence on the strength of the effect of objective numeracy (OR=.64 without control variables vs. OR=.62 with control variables). Consistent with previous findings, the model also revealed that patients who were older and those who perceived their symptoms to be less severe at onset tended to have longer decision times. Actually, the effect of numeracy was of a comparable size to that of perceived severity of symptoms. Detailed results are shown in Table 2.

Table 2. Multiple ordinal regression model predicting decision delay. OR=odds ratio. LLCI/ULCI=Lower/Upper level 95% confidence intervals for the odds ratios. Chi^2 = Wald Chi^2.

Predictor	B	OR	LLCI	ULCI	Chi^2	p
Gender [male]	0.22	1.24	0.38	4.04	0.13	.718
Education	-0.05	0.95	0.65	1.39	0.08	.784
Age	**0.05**	**1.05**	**1.00**	**1.11**	4.48	.034
Number of obstructed arteries	-0.04	0.97	0.63	1.48	0.03	.872
Days from CV event to questionnaire	0.02	1.02	0.95	1.10	0.27	.603
Travel delay	0.09	1.09	0.82	1.45	0.37	.544
Perceived severity of symptoms	**-0.47**	**0.63**	**0.49**	**0.81**	13.27	.000
Hospital anxiety and depression	0.04	1.04	0.98	1.09	1.81	.178
Subjective numeracy	-0.01	0.99	0.94	1.05	0.04	.848
Objective numeracy	**-0.48**	**0.62**	**0.40**	**0.95**	4.72	.030

Note: The number of obstructed arteries and number of days elapsed from the CV event to completing the questionnaire control for patients' physical state influencing numeracy and its effect on decision delay. Scores on the hospital anxiety and depression scale control for patients' mental state. Travel delay, which was estimated by patients, controls for patients' tendencies to recall or estimate the passage of time. We also checked whether controlling for CVD history, symptom knowledge, type of myocardial infarction (STEMI vs. non-STEMI), number of doctor visits, or marital status (married vs. not married) influenced the relationship of numeracy to decision delay. However, these variables had trivial influence.

4. Discussion

People who are less numerate are about four times more likely to decide to delay critically needed medical treatment for more than one hour during ACS, profoundly increasing their risk for death and major disability (French, 2000). This result is in line with our hypothesis and it is based on analyses of our retrospective study of survivors of ACS at an urban hospital in Spain. The estimated magnitude of the benefits of numeracy were independent of many other assessed cognitive, social, health, and demographic factors known to influence decision delay such as age and symptom severity. To the extent the current results generalize, model estimates indicate that numeracy or lack thereof may be among the largest decision delay risks yet to be identified in the literature. Even considering the methodological challenges inherent in comparing across studies and samples, the estimated protective benefits of numeracy appears to far exceed other cognitive factors including knowledge of ACS symptoms. Because seeking medical attention quickly during ACS reduces the risks of major complications that lower quality of life and require follow-up care, the current results also suggest that more numerate people are much less likely to cause significant recurring ACS-related burdens for their health systems and for their families (Berger et al., 1999; Goldberg et al., 1992; Goldberg et al., 1998; Newby et al., 1996). The discovery that numeracy is one of the strongest predictors of ACS decision delay has far reaching implications for (1) improving our understanding and investigations of the causes of decision delay, and (2) improving our ability to identify and educate those who are at risk.

4.1. Risk Literacy and Informed Decision Making

Decision making is a skill. Some people consistently make better decisions across very diverse tasks from investing, to social judgments, to health choices, and beyond (Bruine de Bruin, Parker, & Fischhoff, 2007; Cokely et al., 2012; Ghazal, Cokely, & Garcia-Retamero, 2014; Parker &

Fischhoff, 2005; Peters, 2012). For typical individuals, the link between numeracy and skilled decision making tends to be largely independent of intelligence and other general cognitive abilities (Cokely et al., 2012; Dieckmann et al., 2015; Finucane & Gullion, 2010; Ghazal et al., 2014; Liberali, Reyna, Furlan, Stein, & Pardo, 2012). It is unlikely that when experiencing chest pain, stress, and fear, numerate people start calculating probabilities in their heads. Instead of involving special capacities that enable cold, abstract, and "logical" decision analyses, numerate people typically make better decisions because they deliberate while thinking about their thinking (i.e., metacognition), which helps them expose and avoid biases as they evaluate the personally meaningful aspects of the decision and its consequences (Bruine de Bruin et al., 2007; Cokely et al., 2012; Cokely & Kelley, 2009; Garcia-Retamero, Cokely, & Hoffrage, 2015; Ghazal et al., 2014; Parker & Fischhoff, 2005; Peters, 2012; Reyna et al., 2009). Although numerate participants do benefit from their greater facility with numbers, these same informed decision making processes have been found in studies of the benefits of visual aids that help less numerate people understand specific risks and make decisions as well as highly numerate individuals (Garcia-Retamero & Cokely, 2011; Garcia-Retamero et al., 2015; Garcia-Retamero & Cokely, 2013; Garcia-Retamero & Cokely, 2014; Zikmund-Fisher et al., 2008). By developing an integrate emotional and cognitive understanding of risks, people can "feel" the weight of various decisions and consequences, giving rise to a strong and sustained desire to take appropriate action (Petrova, Garcia-Retamero, & Cokely, 2015; Petrova, van der Pligt, & Garcia-Retamero, 2014).

Unfortunately, the design of the current retrospective study does not allow a detailed analysis of decision processes. We speculate that key mechanisms involved deliberation and elaboration on (1) major risks, (2) potential consequences, and (3) courses of action (Cokely & Kelley, 2009). Instead of focusing on the chances that symptoms were serious, numerate patients may have more frankly considered and discussed the possibility that they were experiencing ACS. Visualizing the most important consequences in personally relevant ways would create hard to ignore, vivid mental imagery (e.g., "my family would be devastated"). National sample studies in the U.S. and Germany indicate that numerate participants do strategically ignore the probability of events to focus on avoiding the worst possible outcome when useful (Pachur & Galesic, 2013). Numerate patients might also explore and consider various courses of action, forming specific behavioral intentions ("if this lasts 10 more minutes I'll wake my spouse"), a cognitive strategy that promotes protective actions in high-stakes situations (e.g., sexual health prevention and detection; Dieckmann et al., 2015; Peters, Baker, Dieckmann, Leon, & Collins, 2010).

It is noteworthy that a growing body of research including results from the current study suggest that even small increases in numeracy can provide considerable inoculation against decision biases (Peters et al., 2010). Most patients in our study had very low numeracy and answered no more than 1 question correctly on a relatively easy test of numeracy (e.g., around the 10th percentile for EU samples). More research is needed to test the benefits of numeracy across the full range of skill and estimate potential threshold levels for protective benefits (e.g., is the 20th percentile enough to maximum risk reduction or do the benefits monotonically increase across skill levels (Ghazal et al., 2014)). Future studies should also test the robustness of our estimates with more diverse samples from other countries using different healthcare systems. For example, Spain offers free healthcare and virtually everyone has health insurance, making financial concerns unlikely. However, numeracy can be even more important in the decisions of patients from countries where financial concerns are likely to play a role in the decision to seek care. For example, lack of insurance and financial concerns about treatment have been associated with longer delays in the United States (Smolderen et al., 2010).

4.2. Reducing Decision Biases and Health Risks

One means of reducing decision delay risk is educating those at risk (Ginde et al., 2008). Unfortunately, a 2012 review of major risk education studies showed that 75% failed to produce any significant reductions in relevant risks (Mooney et al., 2012). Given that previous and current findings indicate there is no robust relation between symptom knowledge and delays, these educational failures may largely reflect too much emphasis on promoting symptom knowledge. In contrast, efforts focusing on emotional responses, education about denial or the costs of delay show more promise and accord with the psychologically plausible mechanisms outlined in the previous section (see for example Mooney et al., 2014; O'Brien, McKee, Mooney, O'Donnell, & Moser, 2014). That said, it seems major implications of the current results may have more to do with creating messages that can be "appropriately tailored and targeted at those at greatest risk and at the people on whom they are likely to call when symptoms arise" (Mooney et al., 2012).

Regardless of its mechanisms, numeracy appears to be among the strongest unique predictors of decision delay. This suggests that a numeracy test itself could be used to reach those who are at greater risk. Setting aside practical, legal, and logistical constraints, if a healthcare professional directly asked numeracy questions to people at higher risk for ACS the answers would likely be diagnostic for delay decision risk. Patients who answer just one question wrong (e.g., "if you flip a fair coin 1000 times, how many times do you think it will come back heads?") appear at least twice as likely to delay decisions (OR=2.30, 95% CI 1.12, 4.72, p=.023). More formally, simple and adaptive online decision aids and education could also be created and incorporated in routine healthcare practices (e.g., numeracy tests included during appointment check-in or annual check-ups). In the current study only numeracy, age, and symptom severity were non-redundant predictors, meaning that if people who are not currently experiencing ACS symptoms answered just four questions on a computer (i.e., age and three numeracy questions) a simple program could instantly provide an estimate of the their decision delay risk (e.g., "you are almost 5 times more likely to wait to seek treatment during a heart attack more than an hour, lowering your survival risk by…click here for more information about your risk and what you can do"). Standards for decision aids (Elwyn et al., 2006; Garcia-Retamero & Cokely, 2013), and optimal learning and memory (e.g., active, elaborative processing; retrieval practice; spacing effects) can also be used to increase efficiency and user-friendliness of such systems (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Okan, Garcia-Retamero, Cokely, & Maldonado, 2015).

4.3. Limitations and Future Directions

The design of the current study was retrospective, such that numeracy was assessed after the cardiac event and decision times were based on self-reported patient recall. Although results are consistent with theoretical models, this design does not allow us to draw certain causal conclusions regarding the mechanisms of the relations between numeracy on decision delay. Because patients completed the numeracy tests just days after a major cardiac event, it is also possible that the severity of their disease or their current mental and physical state could have influenced their results on the objective numeracy measure (Ginde et al., 2008). To help control for this potential confound, we measured the mental state of patients and disease severity using several indicators finding that patients with more obstructed arteries had lower objective numeracy scores. This suggests that disease severity was related to numeracy. However, by statistically controlling for this linkage in our analyses there is good reasons to be confident that our estimate on the relations between numerate and delay decision are likely robust and if anything may underestimate the true relationship in the current sample.

Nevertheless, we note that numeracy could in theory have interacted with disease severity, such that patients with low numeracy may develop more severe ACS due to difficulty sticking to medication regimen or healthy lifestyle (Cavanaugh et al., 2008; Estrada et al., 2004; Waldrop-Valverde et al., 2010). To help address this potential concern we also measured subjective numeracy, which should be less influenced by the patient's current health state. The presence of a marginally significant trend with a relatively large effect size in the expected direction provides converging evidence of a strong and unique association between numeracy and reduced delay decisions during ACS. Moreover, although it is theoretically possible that people with higher numeracy are just quicker to seek medical attention in general, even in cases when this is not warranted (e.g., false ACS alarms or minor illness), several studies suggest that higher health literacy and numeracy tend to be associated with fewer emergency care visits (Berkman et al., 2011; McNaughton et al., 2013).

Questionnaires were administered on average 5 days after the CV event when patients had begun to recover from the episode but memories from the event were still fresh. We also made a special effort to re-invite those patients who declined participation due to feeling sick or upset at the moment. However, the generalizability of our results should still be carefully considered in the light of our sample characteristics. Our sample consisted of ACS survivors, which excludes patients who delayed too long or had a too severe ACS to survive the episode. Sadly, about one in three heart attack patients die before reaching the hospital (DeVon et al., 2010; Heart Research Institute UK, 2015). Similarly, exclusion of the most severely ill patients due to inability to answer the questionnaire means that we could not assess the linkages in an especially relevant group. Theoretically, it follows that among these groups numeracy is also likely to confer some protective benefits, but theory and data do not always converge. To some great degree this concern extends to essentially every study of delay decisions in the literature. While we hope future studies will find ways to address this limitation, the data from the current study still speak loud and clear. Numeracy appears to be among the strongest predictors of decision delay ever identified in ACS patients who seek treatment and survive. We strongly encourage others to help verify, refine, and extend this finding and to take advantage of major unexplored opportunities for risk assessment and individualized interventions that could significantly cut the risk of death and major disability caused by ACS treatment delay decisions.

5. References

Abdel-Kader, K., Dew, M. A., Bhatnagar, M., Argyropoulos, C., Karpov, I., Switzer, G., & Unruh, M. L. (2010). Numeracy skills in CKD: Correlates and outcomes. *Clinical Journal of the American Society of Nephrology, 5*(9), 1566-1573.

Amsterdam, E. A., Wenger, N. K., Brindis, R. G., Casey, D. E., Ganiats, T. G., Holmes, D. R., . . . Kontos, M. C. (2014). 2014 AHA/ACC guideline for the management of patients with non–ST-elevation acute coronary syndromes: A report of the American college of Cardiology/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology, 64*(24), e139-e228.

Apter, A. J., Cheng, J., Small, D., Bennett, I. M., Albert, C., Fein, D. G., . . . Van Horne, S. (2006). Asthma numeracy skill and health literacy. *Journal of Asthma, 43*(9), 705-710.

Arrebola Moreno, A., Garcia-Retamero, R., Catena, A., Alvarez, R. M. M., Moreno, R. M., & Ramírez-Hernández, J. A. (2014). On the protective effect of resilience in patients with acute coronary syndrome. *International Journal of Clinical and Health Psychology, 14*(2), 111-119.

Berger, P. B., Ellis, S. G., Holmes, D. R.,Jr, Granger, C. B., Criger, D. A., Betriu, A., . . . Califf, R. M. (1999). Relationship between delay in performing direct coronary angioplasty and early clinical outcome in patients with acute myocardial infarction: Results from the global use of strategies to open occluded arteries in acute coronary syndromes (GUSTO-IIb) trial. *Circulation, 100*(1), 14-20.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine, 155*(2), 97-107.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology, 92*(5), 938-956.

Burnett, R. E., Blumenthal, J. A., Mark, D. B., Leimberger, J. D., & Califf, R. M. (1995). Distinguishing between early and late responders to symptoms of acute myocardial infarction. *The American Journal of Cardiology, 75*(15), 1019-1022.

Cavanaugh, K., Huizinga, M. M., Wallston, K. A., Gebretsadik, T., Shintani, A., Davis, D., . . . Cherrington, A. (2008). Association of numeracy and diabetes control. *Annals of Internal Medicine, 148*(10), 737-746.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*, 25-47.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4*(1), 20-33.

Cokely, E. T., Ghazal, S., & Garcia-Retamero, R. (2014). Measuring numeracy. In B. L. Anderson & J. Schulkin (Eds.), *Numerical Reasoning in Judgments and Decision Making about Health* (pp. 11-38). Cambridge University Press, Cambridge, UK.

DeVon, H. A., Hogan, N., Ochs, A. L., & Shapiro, M. (2010). Time to treatment for acute coronary syndromes: The cost of indecision. *The Journal of Cardiovascular Nursing, 25*(2), 106-114.

Dieckmann, N., Peters, E., Leon, J., Benavides, M., P Baker, D., & Norris, A. (2015). The role of objective numeracy and fluid intelligence in sex-related protective behaviors. *Current HIV Research*, *13*(5), 337-346.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4-58.

Elwyn, G., O'Connor, A., Stacey, D., Volk, R., Edwards, A., Coulter, A., . . . Bernstein, S. (2006). Developing a quality criteria framework for patient decision aids: Online international Delphi consensus process. *BMJ, 333*(7565), 417.

Estrada, C. A., Martin-Hryniewicz, M., Peek, B. T., Collins, C., & Byrd, J. C. (2004). Literacy and numeracy skills and anticoagulation control. *The American Journal of the Medical Sciences, 328*(2), 88-93.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale. *Medical Decision Making, 27*(5), 672-680.

Finucane, M. L., & Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging, 25*(2), 271.

French, W. J. (2000). Trends in acute myocardial infarction management: Use of the national registry of myocardial infarction in quality improvement. *The American Journal of Cardiology, 85*(5), 5-9.

Galesic, M., & Garcia-Retamero, R. (2011). Communicating consequences of risky behaviors: Life expectancy versus risk of disease. *Patient Education and Counseling, 82*(1), 30-35.

Garcia-Retamero, R., Cokely, E. T., & Hoffrage, U. (2015). Visual aids improve diagnostic inferences and metacognitive judgment calibration. *Frontiers in Psychology, 6,* 932.

Garcia-Retamero, R., Cokely, E. T., Wicki, B., & Joeris, A. (2016). Improving risk literacy in surgeons. *Patient Education and Counselin.* Advance online publication.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied, 17*(3), 270-287.

Garcia-Retamero, R., & Galesic, M. (2013). *Transparent communication of health risks: Overcoming cultural differences*. New York: Springer.

Garcia-Retamero, R., Andrade, A., Sharit, J., & Ruiz, J. G. (2015). Is patients' numeracy related to physical and mental health? *Medical Decision Making, 35*(4), 501-511.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science, 22*(5), 392-399.

Garcia-Retamero, R., & Cokely, E. T. (2014). The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making, 27*(2), 179-189.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Ginde, A. A., Clark, S., Goldstein, J. N., & Camargo, C. A. (2008). Demographic disparities in numeracy among emergency department patients:

Evidence from two multicenter studies. *Patient Education and Counseling, 72*(2), 350-356.

Goldberg, R. J., Gurwitz, J., Yarzebski, J., Landon, J., Gore, J. M., Alpert, J. S., . . . Dalen, J. E. (1992). Patient delay and receipt of thrombolytic therapy among patients with acute myocardial infarction from a community-wide perspective. *The American Journal of Cardiology, 70*(4), 421-425.

Goldberg, R. J., Mooradd, M., Gurwitz, J. H., Rogers, W. J., French, W. J., Barron, H. V., & Gore, J. M. (1998). Impact of time to treatment with tissue plasminogen activator on morbidity and mortality following acute myocardial infarction (the second national registry of myocardial infarction). *The American Journal of Cardiology, 82*(3), 259-264.

Heart Research Institute UK. (2015). *UK heart facts.* Retrieved on 28 February 2016 from http://www.hriuk.org/about-heart-disease/facts-about-heart-disease.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25*(4), 361-381.

Mackay, M. H., Ratner, P. A., Nguyen, M., Percy, M., Galdas, P., & Grunau, G. (2014). Inconsistent measurement of acute coronary syndrome patients' pre-hospital delay in research: A review of the literature. *European Journal of Cardiovascular Nursing, 13*(6), 483-493.

Marden, S., Thomas, P., Sheppard, Z., Knott, J., Lueddeke, J., & Kerr, D. (2012). Poor numeracy skills are associated with glycaemic control in type 1 diabetes. *Diabetic Medicine, 29*(5), 662-669.

McNaughton, C. D., Collins, S. P., Kripalani, S., Rothman, R., Self, W. H., Jenkins, C., . . . Storrow, A. B. (2013). Low numeracy is associated with increased odds of 30-day emergency department or hospital recidivism for patients with acute heart failure. *Circulation, 6*(1), 40-46.

Mooney, M., McKee, G., Fealy, G., O'Brien, F., O'Donnell, S., & Moser, D. (2014). A randomized controlled trial to reduce prehospital delay time in patients with acute coronary syndrome (ACS). *The Journal of Emergency Medicine, 46*(4), 495-506.

Mooney, M., McKee, G., Fealy, G., O'Brien, F., O'Donnell, S., & Moser, D. (2012). A review of interventions aimed at reducing pre-hospital delay time in acute coronary syndrome: What has worked and why? *European Journal of Cardiovascular Nursing, 11*(4), 445-453.

Moser, D. K., Kimble, L. P., Alberts, M. J., Alonzo, A., Croft, J. B., Dracup, K., . . . American Heart Association Council on Cardiovascular Nursing and Stroke Council. (2007). Reducing delay in seeking treatment by patients with acute coronary syndrome and stroke: A scientific statement from the American Heart Association Council on cardiovascular nursing and stroke council. *The Journal of Cardiovascular Nursing, 22*(4), 326-343.

Newby, L. K., Rutsch, W. R., Califf, R. M., Simoons, M. L., Aylward, P. E., Armstrong, P. W., . . . Van de Werf, F. (1996). Time from symptom onset to treatment and outcomes after thrombolytic therapy. *Journal of the American College of Cardiology, 27*(7), 1646-1655.

Nguyen, H. L., Saczynski, J. S., Gore, J. M., & Goldberg, R. J. (2010). Age and sex differences in duration of prehospital delay in patients with acute myocardial infarction: A systematic review. *Circulation, 3*(1), 82-92.

O’Brien, F., McKee, G., Mooney, M., O’Donnell, S., & Moser, D. (2014). Improving knowledge, attitudes and beliefs about acute coronary syndrome through an individualized educational intervention: A randomized controlled trial. *Patient Education and Counseling, 96*(2), 179-187.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2015). Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *Journal of Experimental Psychology: Applied*, *21*(2), 178-194.

Osborn, C. Y., Cavanaugh, K., Wallston, K. A., & Rothman, R. L. (2010). Self-efficacy links health literacy and numeracy to glycemic control. *Journal of Health Communication, 15*, 146-158.

Pachur, T., & Galesic, M. (2013). Strategy selection in risky choice: The impact of numeracy, affect, and cross- cultural differences. *Journal of Behavioral Decision Making, 26*(3), 260-271.

Park, H., Yoon, C., Kang, S., Choi, D., Kim, H., Cho, M. C., . . . Gwon, H. (2013). Early-and late-term clinical outcome and their predictors in patients with ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction. *International Journal of Cardiology, 169*(4), 254-261.

Parker, A. M., & Fischhoff, B. (2005). Decision- making competence: External validation through an individual- differences approach. *Journal of Behavioral Decision Making, 18*(1), 1-27.

Perkins-Porras, L., Whitehead, D. L., Strike, P. C., & Steptoe, A. (2009). Pre-hospital delay in patients with acute coronary syndrome: Factors associated with patient decision time and home-to-hospital delay. *European Journal of Cardiovascular Nursing, 8*(1), 26-33.

Peters, E., & Bjalkebring, P. (2014). Multiple numeric competencies: When a number is not just a number. *Journal of Personality & Social Psychology*, *108*(5), 802-822.

Peters, E., Baker, D. P., Dieckmann, N. F., Leon, J., & Collins, J. (2010). Explaining the effect of education on health: A field study in Ghana. *Psychological Science, 21*(10), 1369-1376.

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs, 26*(3), 741-748.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31-35.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*(5), 407-413.

Petrova, D., Garcia-Retamero, R., & Cokely, E. T. (2015). Understanding the harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making, 35*(7), 847-858.

Petrova, D. G., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making, 27*, 191-199.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment & Decision Making, 7*(3), 332-359.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135*(6), 943-973.

Riegel, B., McKinley, S., Moser, D. K., Meischke, H., Doering, L., & Dracup, K. (2007). Psychometric evaluation of the acute coronary syndrome (ACS) response index. *Research in Nursing and Health, 30*(6), 584-594.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*(11), 966-972.

Smolderen, K. G., Spertus, J. A., Nallamothu, B. K., Krumholz, H. M., Tang, F., Ross, J. S., . . . Chan, P. S. (2010). Health care insurance, financial concerns in accessing care, and delays to hospital presentation in acute myocardial infarction. *JAMA, 303*(14), 1392-1400.

Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., & Brand, H. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health, 12*(1), 80.

Turpie, A. G. (2006). Burden of disease: Medical and economic impact of acute coronary syndromes. *American Journal of Managed Care, 12*(16), S430.

Waldrop-Valverde, D., Jones, D. L., Gould, F., Kumar, M., & Ownby, R. L. (2010). Neurocognition, health-related reading literacy, and numeracy in medication management for HIV infection. *AIDS Patient Care and STDs, 24*(8), 477-484.

World Health Organization. (2014). *The top 10 causes of death. Fact sheet N°310.* Retrieved on 28 February 2016 from http://www.who.int/mediacentre/factsheets/fs310/en/.

Yee, L. M., & Simon, M. A. (2014). The role of health literacy and numeracy in contraceptive decision-making for urban Chicago women. *Journal of Community Health, 39*(2), 394-399.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making, 27*(5), 663-671.

Zikmund-Fisher, B. J., Ubel, P. A., Smith, D. M., Derry, H. A., McClure, J. B., Stark, A., . . . Fagerlin, A. (2008). Communicating side effect risks in a tamoxifen prophylaxis decision aid: The debiasing influence of pictographs. *Patient Education and Counseling, 73*(2), 209-214.

CHAPTER 7:
Emotional distress predicts disease severity

The content of this chapter will be published as Garcia-Retamero, R., Petrova, D., Arrebola Moreno, A., Catena, A., & Ramírez Hernández, J. A. Type D personality is related to severity of acute coronary syndrome. *British Journal of Health Psychology.*

Type D personality is related to severity of acute coronary syndrome

Our aim was to investigate the relationship between Type D (distressed) personality and cardiac biomarkers of disease severity in patients with acute coronary syndrome, and to identify potential mechanisms behind the effect of Type D personality on cardiovascular disease (CVD). Patients (N=215) with acute coronary syndrome completed a survey including a measure of Type D personality. Fasting blood samples including a lipid profile and cardiac enzymes were taken within three days after the cardiovascular event. Data were analyzed using simple correlations, multiple regressions, and mediation analyses. Type D personality was more predictive of severity of the acute coronary syndrome among patients with previous CVD compared to patients without CVD history. Among patients with a previous CVD, Type D personality was associated with ST-elevation and more damage to the myocardium as indicated by higher troponin-I and myoglobin levels. These effects were independent from demographics, CV risk factors, and depression. Lower HDL cholesterol levels mediated the relationship between Type D personality and disease severity. Type D personality was related to a worse lipid profile and more severe acute coronary syndrome in patients with previous history of CVD. These results suggest that severity of the myocardial infarction is a potential mechanism explaining increased mortality in Type D patients with recurrent CVD.

1. Introduction

Type D or distressed personality has been described as the stable tendency to experience simultaneously high negative affectivity and high social inhibition (Denollet, 1998; Denollet, 2005). People with Type D personality often feel negative emotions and inhibit the expression of these emotions in social interactions (Denollet, 1998). Type D patients are irritable, tend to worry, and take a gloomy view of life, but do not share these emotions with other people because of fear of rejection and disapproval (Pedersen & Denollet, 2003). The synergistic effect of negative affectivity and social inhibition has been associated with poor health-related quality of life (Aquarius, Denollet, Hamming, Henegouwen, Van Berge, & De Vries, 2007; Pedersen et al., 2006; Pedersen, Theuns, Muskens-Heemskerk, Erdman, & Jordaens, 2007; Schiffer, Pedersen, Widdershoven, & Denollet, 2008; Schiffer et al., 2005) and poor prognosis in cardiac patients (Denollet, Vaes, & Brutsaert, 2000; Denollet et al., 2006).

A meta-analysis conducted by O'Dell, Masters, Spielmans, and Maisto (2011) concluded that Type D personality is an independent predictor of poor health status and major adverse cardiac events (e.g., cardiac death) in patients with a variety of cardiovascular diseases (CVD), with an associated risk similar to that of traditional cardiovascular (CV) risk factors (Martens, Mols, Burg, & Denollet, 2010; O'Dell et al., 2011). In contrast, a more recent meta-analysis conducted by Grande, Romppel, and Barth (2012) concluded that previous studies overestimated the effect of Type D personality on mortality and cardiac events. For instance, several recent studies showed no effect of Type D personality on mortality or major adverse cardiac events (Coyne et al., 2011; Grande et al., 2011; Pelle et al., 2010; Volz et al., 2011).

These issues speak to the need for further research on the effects of Type D personality on cardiovascular health. For example, discrepancies in results can be clarified by investigating the mechanisms through which Type D personality may influence cardiac prognosis and whether such mechanisms vary in different subgroups of patients. The current research aimed to shed light on such potential mechanisms and had the following two aims: (1) to investigate the relationship between Type D personality and disease severity in acute coronary syndrome (ACS) patients, and (2) to investigate the strength of this relationship in patients with vs. without previous CVD history.

1.1. Type D personality and disease severity

Recent research has identified several behavioral and biological mechanisms that may explain the increased risk in Type D cardiac patients (see Denollet, Schiffer, and Spek (2010) for a review). Behavioral mechanisms include not seeking medical care, insufficient adherence to medical treatment, and unhealthy behaviors (Pelle, Schiffer, Smith, Widdershoven, & Denollet, 2010; Schiffer, Denollet, Widdershoven, Hendriks, & Smith, 2007; Williams et al., 2008; Williams, O'Connor, Grubb, & O'Carroll, 2011a). To illustrate, Type D personality has been associated with unhealthy diet, more frequent smoking, less frequent exercise, fewer regular medical checkups, and failure to consult for cardiac symptoms (Gilmour & Williams, 2012; Mommersteeg, Kupper, & Denollet, 2010; Pedersen et al., 2004; Pelle et al., 2010; Schiffer et al., 2007; Svansdottir et al., 2013; Williams et al., 2008; Williams, Abbott, & Kerr, 2015). Biological mechanisms include CV stress reactivity and neuroendocrine and immunological pathways (Conraads et al., 2006; Sher, 2005; van den Broek et al., 2009; von Känel et al., 2009; Whitehead, Perkins-Porras, Strike, Magid, & Steptoe, 2007). For example, Type D patients show heightened physiological reactions to acute stress (Williams, O'Carroll, & O'Connor, 2009).

In contrast, several studies ruled out another potential mechanism: disease severity (Conraads et al., 2006; Denollet & Brutsaert, 1998; Martens et al., 2010; Pedersen & Denollet, 2006; Svansdottir et al., 2012). In particular, these studies showed that Type D personality is not related to disease severity as indicated by patients' left ventricular ejection fraction (Denollet & Brutsaert, 1998) or their number of obstructed arteries (Svansdottir et al., 2012). Interestingly, these studies showed that Type D personality is related to cardiac prognosis even after controlling for disease severity (Denollet et al., 2000; Martens et al., 2010). Researchers concluded that these results show that (1) the Type D personality assessment is not confounded by disease severity, such that patients who experience a more severe episode report more distress, and (2) disease severity is not among the mechanisms that explain the effect of Type D personality on cardiac outcomes (Denollet & Brutsaert, 1998; Svansdottir et al., 2012).

To the best of our knowledge only few studies have related Type D personality to other measures of disease severity. Cardiac biomarkers such as brain natriuretic peptide (BNP) and high-sensitivity C-reactive protein (hsCRP) might be more sensitive measures. However, recent research in CVD patients showed that levels of these biomarkers in patients with Type D personality are similar to those of patients without Type D personality (Coyne et al., 2011; Mommersteeg et al., 2012; Pelle et al., 2009). Currently, troponin-I is considered the biomarker with the highest diagnostic and prognostic value (Adams et al., 1993; Antman et al., 1996; Hamm et al., 2011). However, no study has investigated whether Type D personality predicts levels of cardiac troponin-I in patients with ACS. To fill this gap, in the current study we investigated the relationship between Type D personality and disease severity as indicated by troponin-I. We also measured myoglobin, a complementary early marker of acute myocardial infarction (de Winter, Koster, Sturk, & Sanders, 1995; Mair et al., 1995).

1.2. Type D and previous history of CVD

Another goal of the current research was to investigate whether a previous history of CVD influences the effect of Type D personality on disease severity. This factor might be important because patients with previous history are at higher risk of having another CV event (Smith et al., 2006). Research has also shown that Type D individuals are more prone to ACS recurrence (Martens et al., 2010). The clustering of these risk factors may result in greater predictive power of Type D personality, such that Type D personality may be related to disease severity more strongly or only among patients experiencing recurrent (as opposed to first time) ACS. In fact, Type D personality is a chronic risk factor that may promote the development of episodic risk factors like vital exhaustion or depression that have been related to the severity of the coronary artery disease (Kop, 1997; Pedersen & Denollet, 2003). Symptoms of vital exhaustion and fatigue are more characteristic of Type D vs. non-Type D patients (Pedersen & Middel, 2001). This can result in accumulated risk that may manifest as increased severity of following cardiac events. Type D personality may then be especially predictive of severity of a recurrent (as opposed to a first episode) CVD.

This hypothesis is further supported by research showing that Type D patients may have difficulty following cardiovascular rehabilitation recommendations, thus exposing themselves to risks of more severe subsequent disease. Myocardial infarction is a frightening event which requires difficult psychological and lifestyle adjustments. Adherence to medication and healthy lifestyle recommendations following an ACS episode are vital to the success of cardiovascular rehabilitation and reducing the risk of recurrence of ACS (Perk et al., 2012). However, research has shown that compared to non-Type D patients, Type D patients are less likely to adhere to medical treatment and healthy habits after myocardial infarction (Molloy et al., 2012; Williams et al., 2011a). Type D patients in cardiac rehabilitation also show more maladaptive coping styles, with potential negative consequences for their recovery (Shanmugasegaram et al., 2014; see Polman, Borkoles, & Nicholls, 2010 for similar results in a healthy population). Type D patients are more concerned about their illness, they believe that it might have more serious consequences, expect it to last longer, and perceive it as less controllable (Williams, O'Connor, Grubb, & O'Carroll, 2011b). Finally, compared to non-Type D patients, Type D patients report poorer health status after cardiac rehabilitation (Pelle et al., 2008). All these findings highlighting the potential role of Type D personality in cardiac rehabilitation suggest that Type D personality may be especially predictive of prognosis among patients with a previous CVD history.

In sum, we investigated the relationship between Type D personality and disease severity in patients with ACS. We used several indicators of CVD severity, including the presence of ST-elevation, number of obstructed arteries and stent implants, and peak levels of cardiac enzymes (i.e., troponin-I and myoglobin). We also investigated plausible pathways explaining the relationship between Type D personality and disease severity. In particular, we investigated whether this relationship is influenced by several classical risk factors (Rosengren et al., 2004) including depression, smoking, hypertension, diabetes, body-mass index, and lipid profile. Finally, we investigated the predictive power of Type D personality in patients without a previous history of CVD and in patients with a previous history of CVD, including previous myocardial infarction or ischemic heart disease.

2. Method

2.1. Participants and procedure

Participants were 215 consecutive patients (average age=61 years, SD=10, range from 31 to 82; 82% male) who were admitted to a hospital with ACS. ACS was defined as a condition compatible with acute myocardial ischemia and/or infarction due to an abrupt reduction in coronary blood flow (Amsterdam et al., 2014). Accordingly, the inclusion criteria for the study were having elevated cardiac markers (i.e., Troponin-I) and clinical and electrocardiographic signs of ischemia. The exclusion criteria were having an inflammatory disease and/or having neurological problems. Participants were recruited between June 2009 and June 2013, approximately 4 days after the ACS. Patients were recruited by a trained assistant or a physician, who described the purpose of the study and answered questions about the study. Of the approached patients 65% agreed to participate in the study. Those who refused mentioned one or more of the following reasons: lack of time, respondent burden, and/or lack of interest in research.

The Ethics Committee of the hospital approved the methodology of the study. Participants read a brief description of the study and signed the informed consent form. Participants then completed a paper-and-pencil survey. The survey contained standard demographic questions (e.g., age, gender, and education) and assessment of psychological constructs (e.g., Type D personality, anxiety, depression, and stressful life events). Participants' medical history was obtained from the hospital medical records. A fasting venous sample was collected in all patients on hospital admission.

2.2. Measures

2.2.1. Psychological risk factors

Type D personality. Participants completed a validated version of the DS14 questionnaire (Denollet, 1998; Denollet, 2005; Montero, Rueda, & Bermúdez, 2012). This questionnaire consists of 14 items that measure the two components of Type D personality: negative affectivity (7 items, e.g., I am often in a bad mood) and social inhibition (7 items, e.g., I often talk to strangers) on scales from 0 (false) to 4 (true). We obtained a score for each component by summing the corresponding items (range 0-28 for each component).

Hospital Anxiety and Depression Scale. The scale consists of 14 questions that measure anxiety (7 items, e.g., "Worrying thoughts go through my mind.") and depression (7 items, e.g., "I still enjoy the things I used to enjoy.") using answer scales from 0 to 4 (Zigmond & Snaith, 1983). We obtained a score for each construct by summing the corresponding items (range 0-21 for each construct, where a score>10 indicates an abnormal case).

Holmes and Rahe stress scale. The scale includes a list of 43 stressful life events (e.g., death of a spouse) that can contribute to illness (Holmes & Rahe, 1967). Participants indicate if they have experienced each event during their lifetime. Each experienced event is then assigned a "life change" score in accordance with how traumatic the event is. The scores are then summed to arrive at a final stress score (range 0-1467, where a score>300 indicates a high risk of developing a stress-related illness).

2.2.2. CV risk factors.

The following information on classical risk factors, cardiovascular disease (CVD) history, and disease severity was obtained from patients' medical records:

CVD history. Patients had a history of ischemic heart disease or a previous myocardial infarction.

Smoking. If patients smoked daily, they were classified as smokers. Otherwise, they were classified as non-smokers.

Hypertension. Patients were classified as hypertensive if they (a) had an average systolic blood pressure of>140 mm Hg and/or a diastolic blood pressure of >90 mm Hg at rest, (b) had previous history of hypertension, or (c) were taking antihypertensive drugs.

Diabetes. Patients were classified as diabetics if they (a) had fasting blood glucose of >126 mg/dL, (b) had previous history of diabetes mellitus, or (c) were taking insulin/oral hypoglycemics.

Body-mass index (BMI). We computed patients' body-mass index as weight (kg)/ height (m) squared.

Lipid profile. Patients' total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, and triglyceride levels were measured using standard methods.

2.2.3. Indicators of ACS severity

STEMI. We recorded whether participants suffered an ST-segment elevation myocardial infarction (STEMI) or a non-STEMI. STEMI reflects an acute coronary occlusion and indicates a poor early-term clinical outcome. In contrast, a non-STEMI usually reflects a critically ill but not complete coronary artery obstruction and indicates a poor long-term clinical outcome (Park et al., 2013). Patients were classified according to the following criteria (O'Gara et al., 2013): characteristic symptoms of myocardial ischemia in association with persistent electrocardiographic (ECG) ST elevation and subsequent release of biomarkers of myocardial necrosis were classified as STEMI; the absence of persistent ST elevation in combination with elevated cardiac biomarkers in an appropriate clinical context was classified as non-STEMI.

Number of obstructed arteries. We recorded the number of obstructed vessels for each patient. A coronary stenosis of more than 50% in a main branch (i.e., left main, left anterior descending, circumflex, or right coronary artery) was considered an obstructed vessel.

Number of stents. We recorded the number of coronary stents that the patient had implanted during a percutaneous coronary intervention (PCI). In case of previous PCIs, we only considered stents that were implanted as a consequence of the current cardiac episode.

Cardiac biomarkers. Standard enzymatic methods were used to measure peak levels of the cardiac enzymes troponin-I and myoglobin after the coronary event. These enzymes are released after myocardial infarction and serve as indicators of damage to the myocardium (Thygesen et al., 2012).

2.3. Statistical Analyses

Patients' risk factors were analyzed using descriptive statistics. We computed Spearman's Rho rank correlation coefficients to investigate if Type D personality was related to CV risk factors and ACS severity indicators in all patients. We then divided the whole sample in two subsamples: patients with and without CVD history, and investigated the relationship between Type D personality and CV risk factors and ACS severity indicators in patients with and without previous CVD. We also conducted multiple regression analyses to test if Type D personality was related to severity indicators controlling for the effect of demographics, CV risk factors, anxiety, depression, and stress. We used the glm function in R (http://www.r-project.org/) to conduct these analyses. Anxiety, depression, and stress were used as continuous variables. The dependent variables, levels of troponin-I and myoglobin, were highly skewed and were log and square root transformed, respectively, to improve normality. We conducted linear regressions for continuous outcomes (i.e., HDL, troponin-I, and myoglobin), logistic regressions for dichotomous outcomes (i.e., diabetes and STEMI), and Poisson regressions for count outcomes (i.e., obstructed arteries and stent implants).

We conducted two sets of multiple regression analyses: controlling for demographics and CV risk factors (Set 1), and controlling for psychological factors (Set 2). In all multiple regressions, we first included only the control variables (Model 1). In Set 1, the control variables were demographics and those CV risk factors that were associated to the outcome using a liberal cut-off of $p<.1$. In Set 2, the control variables were anxiety, depression, and stressful life events. After including the control variables, we added the Type D personality as an additional predictor (Model 2). We used two criteria to evaluate if Type D personality was a significant predictor. Namely, we tested for significance ($p<.05$) and compared the Akaike Information Criteria (AIC) of Model 1 and Model 2 with smaller AIC indicating a better model such that ΔAIC (Model2−Model1)≥ -1. AIC is a measure of model fit that takes into account both the statistical goodness of fit and the number of estimated parameters and imposes a penalty for increasing the number of parameters (Burnham & Anderson, 2002). We used the percentage of explained variance (R2) as an additional measure of model fit.

Finally, we conducted mediation analyses to investigate whether CV risk factors that were related to Type D personality mediated the effect of Type D personality on CVD severity. We used the PROCESS Macro for SPSS following Preacher and Hayes suitable for small samples (Preacher & Hayes, 2008). We tested for indirect effects of Type D personality on CVD severity by using an accelerated and bias corrected bootstrap procedure based on 5000 samples. We considered an indirect effect significant if the 95% bootstrap confidence intervals for the effect did not include 0. In this case, the percentage of explained variance (R2) was used as measure of model fit.

3. Results

3.1. Participant risk profiles

Table 1 shows descriptive statistics of the continuous variables and recommended values where applicable. On average, patients were overweight, and had higher than normal cardiac biomarker values (i.e., myoglobin and troponin-I), a result that is common in patients with ACS. On average, patients also had lower HDL cholesterol than recommended. However, the averages for the rest of the lipids were within the norm (see Table 1). Thirty-seven percent of the participants were smokers, 59% had hypertension, 36% had diabetes, and 45% had a STEMI. Sixty-five (30%) patients had history of CVD, of which 8% had suffered a previous myocardial infarction, 22% had history of ischemic heart disease, and 70% had history of both. The proportion of Type D patients among those without CVD history (33%) and with CVD history (32%) was similar, Pearson Chi-square=.022, p=.883. The proportions of patients showing abnormally high values on the psychological assessments was 44% for anxiety (>10), 14% for depression (>10), and 85% for stressful life events (>300).

3.2. Type D personality

The two components of Type D personality (i.e., negative affectivity (NA) and social inhibition (SI)) showed good internal consistency (Cronbach α =.83,

.81, .80, for the whole scale, NA, and SI, respectively). In our sample, NA had a mean score of 13.53 (SD=5.90) and a median score of 14 (range 1-25). SI had a mean score of 9.78 (SD=5.08) and a median score of 10 (range 0-25). We used the traditional cut-off approach (i.e., cut-off score of 10 for every component (Emons, Meijer, & Denollet, 2007) to categorize patients into Type D and non-Type D. Following this procedure, we classified 33% of the patients as Type D personality patients.

Table 1. Descriptive statistics for continuous indicators (N=215).Values in parenthesis indicate normal/recommended range.

	Min.	***Max.***	***Mean***	***LLCI***	***ULCI***	***SD***	***Median***
BMI (18.5-24.9 kg/m^2)	20	48	29	28	30	4	28
Anxiety (0-10)	4	17	10.01	9.63	10.40	2.76	10
Depression (0-10)	2	17	8.35	8.03	8.68	2.35	8
Stressful life events (<300)	34	1331	538	507	570	231	515
Stent implants	0	7	1.29	1.13	1.45	1.19	1
Obstructed arteries	0	3	1.77	1.64	1.89	0.93	2
Myoglobin (14-66 ng/ml)	16	4030	464	346	582	700	189
Troponin-I (0,01-0,05ng/ml)	0.01	102	30	24	36	37	15
HDL (40-71 mg/dl)	3	63	38	37	40	10	38
LDL (60-150 mg/dl)	27	219	103	97	109	35	102
Total cholesterol (100-220 mg/dl)	79	347	175	168	183	44	176
Triglycerides (50-230 mg/dl)	53	970	157	141	173	94	135

Note: LLCI=95% Lower level confidence interval. ULCI=95% Upper level confidence interval. *SD*=Standard deviation.

Table 2. Correlations between Type D personality and CV risk factors and ACS severity indicators as a function of CVD history.

	All patients **(N=215)**		Without CVD history **(N=150)**		With CVD history **(N=65)**	
	r	***p***	***r***	***p***	***r***	***p***
Smoking	-.084	.221	-.134	.102	.038	.764
Hypertension	-.039	.569	-.038	.644	-.039	.757
Diabetes	.115	.093	**.220***	**.007**	-.086	.494
BMI	.020	.778	.070	.416	-.076	.555
Anxiety	**.182***	**.008**	**.212***	**.010**	.127	.318
Depression	.073	.291	.087	.290	.048	.707
Stressful life events	.030	.670	.101	.225	-.097	.454
HDL	-.079	.262	-.022	.795	**-.274***	**.031**
LDL	.003	.965	-.047	.586	.121	.351
Total cholesterol	.064	.367	.033	.695	.124	.333
Triglycerides	.108	.125	.091	.283	.148	.247
STEMI	.059	.390	-.047	.567	**.323***	**.009**
Stent implants	.129	.059	.151	.065	.084	.508
Obstructed arteries	.016	.815	.051	.534	-.060	.635
Myoglobin	.036	.612	-.046	.588	**.288***	**.022**
Troponin-I	.050	.478	-.014	.870	**.271***	**.030**

Note. * $p<.05$. BMI: body-mass index; HDL: high-density cholesterol, LDL: low-density cholesterol; STEMI: presence of ST-segment elevation.

Table 3. Results from multiple regression models controlling for demographics and risk factors: Unstandardized regression (B) coefficients and p-values for the effect of Type D personality on CV risk factors and ACS severity indicators among patients without CVD history (N=150) and with CVD history (N=65). Akaike Information Criteria (AIC) and R^2 indices of model fit for models with (M2=Models 2) and without (M1=Models 1) Type D personality.

CVD history	Outcome variables	B	p	AIC M1	AIC M2	ΔAIC M2 − M1	R^2 M1	R^2 M2	Δ R^2 M2 − M1	Factors in M1
No (N=150)	**Diabetes**	1.07	.006	178	172	-6	.04	.08	.04	age, gender
	Stent implants	.28	.059	434	433	-1	.01	.02	.01	age, gender
Yes (N=65)	**HDL cholesterol**	-5.84	.017	445	440	-5	.16	.24	.08	age, gender, TC
	STEMI	1.42	.023	80	77	-3	.10	.17	.07	age, gender, HT
	Myoglobin	.36	.038	123	121	-3	.06	.13	.07	age, gender
	Troponin-I	1.36	.088	317	316	-1	.07	.12	.05	age, gender, HT

Note: The control factors included on the model were demographic variables (age and gender) and risk factors (smoking, hypertension (HT), diabetes, BMI, total cholesterol (TC)) that were related ($p<.1$) to the respective outcome variable. Model 2 included Model 1 control factors plus Type D personality. R^2 *for logistic and Poisson regressions is McFadden pseudo* R^2*.*

Table 4. Results from multiple regression models controlling for psychological factors: Unstandardized regression (B) coefficients and p-values for the effect of Type D personality on CV risk factors and ACS severity indicators among patients without CVD history (N=150) and with CVD history (N=65). Akaike Information Criteria (AIC) and R^2 indices of model fit for models with (M2=Models 2) and without (M1=Models 1) Type D personality.

CVD history	Outcome variables	*B*	*p*	AIC M1	AIC M2	ΔAIC M2 − M1	R^2 M1	R^2 M2	Δ R^2 M2 − M1
No (N=150)	**Diabetes**	.96	.019	169	166	-3	.09	.12	.03
	Stent implants	.31	.050	417	415	-2	.05	.06	.01
Yes (N=65)	**HDL cholesterol**	-4.75	.062	430	428	-2	.01	.07	.06
	STEMI	1.26	.039	81	79	-2	.08	.13	.05
	Myoglobin	.39	.033	123	120	-3	.01	.09	.08
	Troponin-I	1.68	.051	309	307	-2	.02	.08	.06

Note: The control factors included in the models were anxiety, depression, and stressful life events. Model 2 included Model 1 control factors plus Type D personality. R^2 *for logistic and Poisson regressions is McFadden pseudo* R^2*.*

3.3. Relationship of Type D personality with CV risk factors and ACS severity indicators

Table 2 shows that across all patients, Type D personality was not predictive of CV risk factors and ACS severity indicators, with the exception of a small marginally significant correlation with the number of stent implants. Importantly, dividing the sample according to the presence of previous CVD revealed that Type D personality was associated to ACS severity indicators: Type D personality patients with previous CVD were more likely to have a STEMI (55% and 22%, respectively for Type D and non-Type D, p=.009), and had more elevated cardiac biomarker levels, as indicated by both higher myoglobin (M=554, SD=649 and M=220, SD=463, for Type D and non-Type D, respectively, p=.022) and troponin-I (M=26, SD=33, and M=11, SD=21, for Type D and non-Type D, respectively p=.035). In contrast, Type D personality was not associated with a more severe ACS in patients without previous CVD.

In addition, among patients with previous CVD, Type D personality was related to a worse lipid profile as indicated by lower HDL cholesterol levels (M=34.40, SD=7.58 and M=39.70, SD=8.99, respectively for Type D and non-Type D, p=.031). However, Type D personality was not related to LDL, total cholesterol, or triglycerides (all p>.1) in patients with previous CVD.

Among patients without CVD history, Type D patients were more likely to have diabetes (50% and 27%, for Type D and non-Type D, respectively, p=.007), had higher anxiety scores (M=10.64, SD=2.58 and M=9.41, SD=2.72, for Type D and non-Type D, respectively, p=.009), and tended to need more stent implants (M=1.52, SD=1.34 and M=1.15, SD=1.04, for Type D and non-Type D, respectively, p=.065). In this group of patients, Type D personality was not significantly related to other CV risk factors.

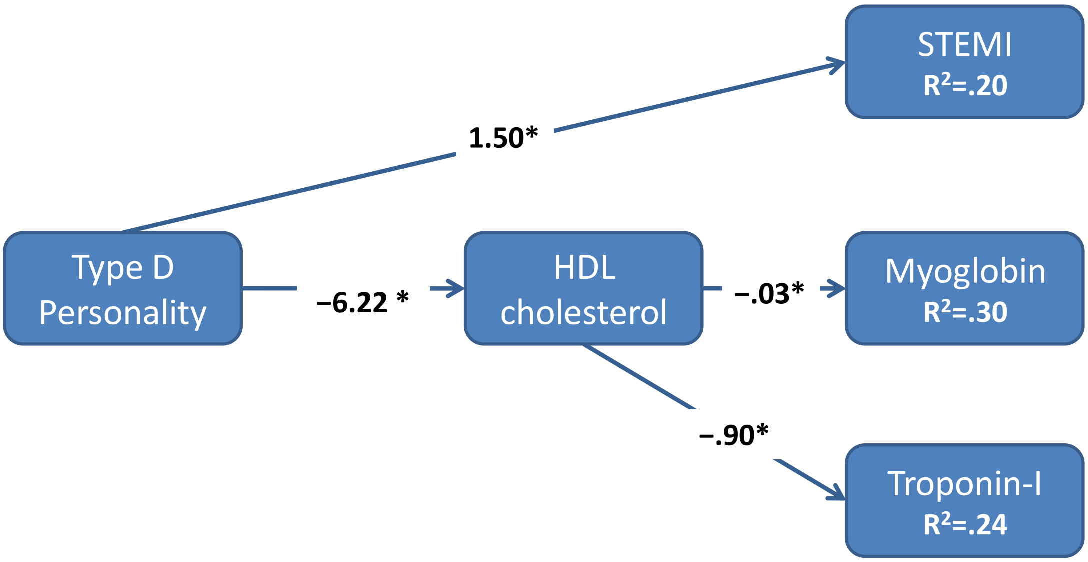
Results in multiple regression analyses showed that the significant first-order associations described above persisted after controlling for basic demographics and risk factors (see Table 3) and psychological factors (see

Table 4). In particular, in all regressions reported in Table 3 we controlled for demographics (age, gender) as well as CV risk factors that were related to the outcome variable. In all regressions reported in Table 4 we controlled for anxiety, depression, and stressful life events. Adding Type D personality to the regression analyses (Models 2) increased model fit for all these outcomes as indicated by a smaller AIC (larger R2) in the model including Type D personality.

3.4. Mediation analysis

We tested whether HDL cholesterol mediated the relationship between Type D personality and STEMI, myoglobin, and troponin-I in patients with a previous CVD. We included HDL cholesterol as a candidate mediator because it was a risk factor significantly associated with Type D personality, and because a potential mediation effect is in line with theories and evidence about biological and behavioral mechanisms of Type D personality (e.g., Kupper & Denollet, 2007; Mommersteeg et al., 2010). We conducted mediation analyses using bootstrapping and controlled for the effect of basic demographics and risk factors (i.e., age, gender, BMI, hypertension, smoking, and total cholesterol). There was no indirect effect of Type D personality on STEMI (Effect=.30, 95% CI [-.53, 1.70]). Rather, after including HDL in the regression, there was a significant direct effect of Type D personality on STEMI, B=1.50, p=.049. In contrast, HDL mediated the relationship between Type D personality and myoglobin (Effect=.16, 95% CI [.02, .39]) and troponin-I levels (Effect=.56, 95% CI [.02, 1.56]) (see Figure 1).

Figure 1. Summary of the three mediation analyses. Direct and indirect effects of Type D personality on each of three ACS severity indicators. Coefficients are unstandardized β. Only significant paths are displayed. *p<.05. R^2=percentage of variance explained.



4. Discussion

The current research extends the literature on the effect of Type D personality in CV health in several ways. To the best of our knowledge, this is the first study investigating the effect of Type D personality on severity of ACS as indicated by the cardiac enzymes troponin-I and myoglobin, and as a function of previous CVD history. Among patients who suffered a first episode of CVD, Type D personality was only weakly related to ACS severity, as indicated by a small correlation between Type D personality and only one of the severity indicators (i.e., number of stent implants). In contrast, among patients who experienced a recurrent CVD (e.g., a second myocardial infarction), Type D personality was related to ACS severity, as indicated by stronger correlations between Type D personality and three severity indicators. In particular, patients with Type D personality were more likely to experience an ACS characterized by ST-segment elevation, which reflects an acute coronary occlusion and indicates a poor early-term clinical outcome. They also suffered a larger damage of the myocardium, as indicated by peak levels of myoglobin and troponin-I. Importantly, the effects of Type D personality were independent from demographics, CV risk factors, and psychological factors like anxiety, depression, and stress.

In contrast to the results reported here, several studies suggested that Type D personality might not be related to cardiac disease severity as indicated by the number of obstructed arteries (Martens et al., 2010; Pedersen & Denollet, 2006; Svansdottir et al., 2012), left ventricular ejection fraction levels (Denollet & Brutsaert, 1998), BNP (Coyne et al., 2001; Pelle, van den Broek, Szabo, & Kupper, 2009), or hsCRP (Mommersteeg, Pelle, Ramakers, Szavo, Denollet, & Kupper, 2012). These results have led researchers to conclude that disease severity was not one of the mechanisms explaining the effect of Type D personality on cardiac prognosis. In contrast, our study suggests that Type D personality can affect cardiac prognosis through disease severity, at least among patients who have recurrent CVD. It is possible that some of the disease severity measures used in previous research (e.g., ejection fraction levels or the number of affected arteries) were not sensitive enough to detect an association with the personality assessment (Martens et al., 2010; Pedersen & Denollet, 2006; Svansdottir et al., 2012). This was also the case in the current study (i.e., there was no correlation between Type D and the number of obstructed arteries). In contrast, the cardiac enzymes used in the current research are sensitive measures of damage to the myocardium and are used as reliable indicators of ACS severity in clinical practice (Thygesen et al., 2012; Hamm et al., 2011). Troponin-I, in particular, is currently the gold standard in diagnosis, prognosis, and management of ACS (del Val Martin, Fernández, & Gómez, 2015; Hamm et al., 2011). While other authors have found no relationship between Type D personality and troponin levels in healthy individuals (Beutel et al., 2012), in ACS patients troponin levels have been associated with resilience – a protective psychological factor (Arrebola Moreno et al., 2014). There are other studies that used highly sensitive and prognostic biomarkers like BNP (Coyne et al., 2011; Pelle et al., 2009) and hsCRP (Mommersteeg et al., 2012) but they found no association with Type D personality. Again, these studies did not examine the predictive power of Type D personality as a function of CVD history.

Overall, a lack of effect of Type D personality on disease severity seems unlikely given the proposed behavioral and biological mechanisms by which Type D personality is assumed to affect CVD etiology and prognosis. For example, as discussed previously, Type D personality individuals are more likely to have an unhealthier lifestyle, including a poor exercise regiment, unhealthy diet, and smoking (Mommersteeg et al., 2010; Williams et al., 2008; 2015). Such factors are associated with a worse lipid profile (Panel NCEPNE, 2002), and potentially a more severe ACS and a higher likelihood of death following the ACS event (Perk et al., 2012).

Given the lack of association between Type D personality and (1) history of CVD and (2) ACS severity in patients without past CVD history, our results speak mostly of the prognostic rather than etiological role of Type D in CVD. Our results suggest that Type D personality may be especially predictive of prognosis and disease severity when clustered with other risk factors (e.g., previous history of myocardial infarction). Much like the way negative affectivity and social inhibition are more predictive as a synergy, Type D may be more predictive in combination with other factors. Such clustering of psychosocial risk factors can substantially and disproportionately increase cardiovascular risk for patients and medical costs for health systems (Rozanski, Blumenthal, & Kaplan, 1999). Identifying subgroups of patients who may be at risk for clustering and intervening to prevent it may be one way to reduce this risk (Pedersen & Denollet, 2003).

The current research extends the literature on the effect of Type D personality by investigating bio-behavioral pathways that can help explain the increase in risk in Type D patients. Mediation analyses showed that the effect of Type D personality on ACS severity among patients with CVD history was accounted for, at least partially, by lower levels in HDL cholesterol. HDL cholesterol has important protective anti-inflammatory and antithrombotic functions (Cesari et al., 2010) and is predictive of CVD morbidity and mortality (Panel NCEPNE, 2002). There are several potential mechanisms through which Type D personality may be affecting HDL levels.

First, to the extent that HDL levels can be controlled by lifestyle changes, these results suggest that Type D personality patients who had already suffered a previous CVD episode may have had less success with cardiac rehabilitation. For example, they could have led more unhealthy lifestyles than non-Type D patients (e.g., less exercise and failure to follow omega-3-rich diet), resulting in a worse lipid profile. Indeed, a number of studies in healthy populations speak in favor of this hypothesis. These studies showed that Type D personality is associated with a less diverse diet, a more sedentary lifestyle, and worse physical fitness (Einvik et al., 2011; Mommersteeg et al., 2010; Svansdottir et al., 2013).Therefore, the increase in risk in recurrent ACS patients with Type D personality could be, at least partially, due to common preventable risk behaviors. Unfortunately, in our study we did not measure exercise or diet, so we could not extend our model to include these variables.

Second, another unmeasured factor that has been related to both Type D personality and success in cardiac rehabilitation is social support. Social support contributes to both primary and secondary prevention of CVD (e.g., Barth, Schneider, & von Känel, 2010; Petrova, Garcia-Retamero, & Catena,

2015) and regular participation in cardiac rehabilitation (e.g., Jackson, Leclerc, Erskine, & Linden, 2005). However, Type D patients often lack social support (Polman et al., 2010; Williams et al., 2008), which may indirectly contribute to failure to adopt a healthier lifestyle and control lipids.

Finally, another mechanism that may explain lower HDL levels in Type D patients is heart rate variability (HRV). HRV is an index of vagal nerve activity and an important marker of emotion regulation ability (Appelhans & Luecken, 2006). Lower high frequency HRV, a marker of worse health outcomes, has been associated with the presence of Type D personality in Caucasians (Martin et al., 2010). Lower HRV tends to be associated with lower HDL levels (Min, Min, Paek, & Cho, 2008), although this relationship has not been consistent in the literature (see Stuckey, 2013 for a review). Hence, Type D patients may have lower HRV, reflecting difficulties in emotion regulation, which may contribute to lower HDL levels. However, no research to our knowledge has investigated HRV and HDL cholesterol levels in Type D patients.

Interestingly, Type D personality was related to HDL cholesterol but was not related to LDL cholesterol or total cholesterol levels (see Einvik et al. (2011) for similar results). A plausible explanation of this finding is that due to their CVD history, these patients were taking cholesterol lowering drugs, which mainly target LDL cholesterol. Although we could not control for patients' medication use before the current cardiac episode, it is common practice in Spain to prescribe cholesterol lowering drugs to ACS patients regardless of cholesterol levels (Steg et al., 2013). We believe it is safe to assume that all patients with previous CVD history in our study were prescribed such drugs. It may be informative in future studies to control for the type, dosage, and adherence to cholesterol medication, as these may be related to Type D personality.

Our study leaves open several questions for future research. A limitation of our study was that it was cross-sectional. Type D personality and disease severity were assessed roughly at the same time, so causality cannot be inferred. However, our mediation analysis was theory driven (e.g., Kupper & Denollet, 2007; Mommersteeg et al., 2010) and we believe it can be informative to the literature of potential mechanisms of Type D personality. Our research also adds cross-cultural evidence to the relatively homogeneous Type D personality literature, which until recently was primarily focused on patients in North-Western European samples (O'Dell et al., 2011). Nevertheless future research can investigate the relationship between Type D personality and CVD using a longitudinal design. Another limitation of our study was that the number of patients with CVD history in our sample was relatively small. However, the proportion was similar to that in previous research on Type D personality (30%, Denollet & Brutsaert, 1998). In an effort to have more statistical power to detect differences, we did not use a statistical correction for multiple tests. This means that our criterion of significance was relatively liberal and that some of the findings may be due to Type-I error. Future studies can replicate our findings in larger samples of patients with previous CVD.

In sum, our research showed that Type D personality is related to a more severe ACS among patients with recurrent CVD. In particular, Type D patients with recurrent CVD had lower levels of HDL cholesterol, which contributed to ACS severity. These results suggest that ACS severity can be a potential mechanism explaining increased mortality in Type D patients with recurrent CVD. More research is needed to clarify the results of Type D personality on CVD severity and long-term cardiovascular outcomes using samples from different countries. Our results suggest that using more sensitive measures (e.g., Troponin-I) and taking into account the presence of previous history of CVD could help to clarify the mixed findings.

5. References

Adams, J. E.,3rd, Bodor, G. S., Davila-Roman, V. G., Delmez, J. A., Apple, F. S., Ladenson, J. H., et al. (1993). Cardiac troponin I. A marker with high specificity for cardiac injury. *Circulation*, *88*(1), 101-106.

Amsterdam, E. A., Wenger, N. K., Brindis, R. G., Casey Jr, D. E., Ganiats, T. G., Holmes Jr, D. R., ... & Levine, G. N. (2014). ACC/AHA guideline for the management of patients with non-ST-elevation acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*, *130*, 2354-2394.

Antman, E. M., Tanasijevic, M. J., Thompson, B., Schactman, M., McCabe, C. H., Cannon, C. P., et al. (1996). Cardiac-specific troponin I levels to predict the risk of mortality in patients with acute coronary syndromes. *New England Journal of Medicine*, *335*(18), 1342-1349.

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, *10*(3), 229.

Aquarius, A. E., Denollet, J., Hamming, J. F., Henegouwen, Van Berge, & De Vries, J. (2007). Type-D personality and ankle brachial index as predictors of impaired quality of life and depressive symptoms in peripheral arterial disease. *Archives of Surgery*, *142*(7), 662-667.

Arrebola Moreno, A., Garcia-Retamero, R., Catena Martínez, A., Alvarez, R. M. M., Moreno, R. M., & Hernández, J. A. R. (2014). On the protective effect of resilience in patients with acute coronary syndrome. *International Journal of Clinical and Health Psychology*, *14*(2), 111-119.

Barth, J., Schneider, S., & von Känel, R. (2010). Lack of social support in the etiology and the prognosis of coronary heart disease: a systematic review and meta-analysis. *Psychosomatic Medicine*, *72*(3), 229-238.

Bassand, J. P., Hamm, C. W., Ardissino, D., Boersma, E., Budaj, A., Fernández-Avilés, F., ... & Wijns, W. (2007). Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes. *European Heart Journal*, *28*(13), 1598-1660.

Beutel, M. E., Wiltink, J., Till, Y., Wild, P. S., Münzel, T., Ojeda, F. M., ... & Zwiener, I. (2012). Type D personality as a cardiovascular risk marker in the general population: results from the Gutenberg health study. *Psychotherapy and Psychosomatics*, *81*(2), 108-117.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer.

Cesari, M., Marzetti, E., Laudisio, A., Antonica, L., Pahor, M., Bernabei, R., et al. (2010). Interaction of HDL cholesterol concentrations on the relationship between physical function and inflammation in community-dwelling older persons. *Age and Ageing*, *39*(1), 74-80.

Conraads, V. M., Denollet, J., De Clerck, L. S., Stevens, W. J., Bridts, C., & Vrints, C. J. (2006). Type D personality is associated with increased levels of tumour necrosis factor (TNF)-α and TNF-α receptors in chronic heart failure. International *Journal of Cardiology*, *113*(1), 34-38.

Coyne, J. C., Jaarsma, T., Luttik, M. L., van Sonderen, E., van Veldhuisen, D. J., & Sanderman, R. (2011). Lack of prognostic value of type D personality for mortality in a large sample of heart failure patients. *Psychosomatic Medicine*, *73*(7), 557-562.

del Val Martin, D., Fernández, M. S., & Gómez, J. L. Z. (2015). Biomarkers in acute coronary syndrome. *IJC Metabolic & Endocrine*, *8*, 20-23.

de Oliveira, E. P., Manda, R. M., Torezan, G. A., Corrente, J. E., & Burini, R. C. (2011). Dietary, anthropometric, and biochemical determinants of plasma high-density lipoprotein-cholesterol in free-living adults. *Cholesterol*, *2011*, 851-750.

de Winter, R. J., Koster, R. W., Sturk, A., & Sanders, G. T. (1995). Value of myoglobin, troponin T, and CK-MBmass in ruling out an acute myocardial infarction in the emergency room. *Circulation*, *92*(12), 3401-3407.

Denollet, J. (1998). Personality and coronary heart disease: The type-D scale-16 (DS16). *Annals of Behavioral Medicine*, *20*(3), 209-215.

Denollet, J. (2005). DS14: Standard assessment of negative affectivity, social inhibition, and type D personality. *Psychosomatic Medicine*, *67*(1), 89-97.

Denollet, J., & Brutsaert, D. L. (1998). Personality, disease severity, and the risk of long-term cardiac events in patients with a decreased ejection fraction after myocardial infarction. *Circulation*, *97*(2), 167-173.

Denollet, J., Pedersen, S. S., Ong, A. T., Erdman, R. A., Serruys, P. W., & van Domburg, R. T. (2006). Social inhibition modulates the effect of negative emotions on cardiac prognosis following percutaneous coronary intervention in the drug-eluting stent era. *European Heart Journal*, *27*(2), 171-177.

Denollet, J., Schiffer, A. A., & Spek, V. (2010). A general propensity to psychological distress affects cardiovascular outcomes: Evidence from research on the type D (distressed) personality profile. *Circulation, 3*(5), 546-557.

Denollet, J., Vaes, J., & Brutsaert, D. L. (2000). Inadequate response to treatment in coronary heart disease: Adverse effects of type D personality and younger age on 5-year prognosis and quality of life. *Circulation*, *102*(6), 630-635.

Einvik, G., Dammen, T., Hrubos-Strom, H., Namtvedt, S. K., Randby, A., Kristiansen, H. A., et al. (2011). Prevalence of cardiovascular risk factors and concentration of C-reactive protein in type D personality persons without cardiovascular disease. *European Journal of Cardiovascular Prevention and Rehabilitation*, *18*(3), 504-509.

Emons, W. H., Meijer, R. R., & Denollet, J. (2007). Negative affectivity and social inhibition in cardiovascular disease: Evaluating type-D personality and its assessment using item response theory. *Journal of Psychosomatic Research*, *63*(1), 27-39.

Gilmour, J., & Williams, L. (2012). Type D personality is associated with maladaptive health-related behaviours. *Journal of Health Psychology*, *17*(4), 471-478.

Grande, G., Romppel, M., & Barth, J. (2012). Association between type D personality and prognosis in patients with cardiovascular diseases: A systematic review and meta-analysis. *Annals of Behavioral Medicine*, *43*(3), 299-310.

Grande, G., Romppel, M., Vesper, J. M., Schubmann, R., Glaesmer, H., & Herrmann-Lingen, C. (2011). Type D personality and all-cause mortality in cardiac patients--data from a german cohort study. *Psychosomatic Medicine*, *73*(7), 548-556.

Hamm, C. W., Bassand, J. P., Agewall, S., Bax, J., Boersma, E., Bueno, H., ... & Ohman, M. (2011). ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *European Heart Journal*, *32*(23), 2999-3054.

Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research*, *11*(2), 213-218.

Jackson, L., Leclerc, J., Erskine, Y., & Linden, W. (2005). Getting the most out of cardiac rehabilitation: a review of referral and adherence predictors. *Heart*, *91*(1), 10-14.

Kop, W. J. (1997). Acute and chronic psychological risk factors for coronary syndromes: moderating effects of coronary artery disease severity. *Journal of Psychosomatic Research*, *43*(2), 167-181.

Kupper, N., & Denollet, J. (2007). Type D personality as a prognostic factor in heart disease: assessment and mediating mechanisms. *Journal of Personality Assessment*, *89*(3), 265-276.

Mair, J., Morandell, D., Genser, N., Lechleitner, P., Dienstl, F., & Puschendorf, B. (1995). Equivalent early sensitivities of myoglobin, creatine kinase MB mass, creatine kinase isoform ratios, and cardiac troponins I and T for acute myocardial infarction. *Clinical Chemistry*, *41*(9), 1266-1272.

Martens, E. J., Mols, F., Burg, M. M., & Denollet, J. (2010). Type D personality predicts clinical events after myocardial infarction, above and beyond disease severity and depression. *The Journal of Clinical Psychiatry*, *71*(6), 778-783.

Martin, L. A., Doster, J. A., Critelli, J. W., Lambert, P. L., Purdum, M., Powers, C., & Prazak, M. (2010). Ethnicity and Type D personality as predictors of heart rate variability. *International Journal of Psychophysiology*, *76*(2), 118-121.

Min, K. B., Min, J. Y., Paek, D., & Cho, S. I. (2008). The Impact of the Components of Metabolic Syndrome on Heart Rate Variability: Using the NCEP-ATP III and IDF Definitions. *Pacing and Clinical Electrophysiology*, *31*(5), 584-591.

Molloy, G. J., Randall, G., Wikman, A., Perkins-Porras, L., Messerli-Burgy, N., & Steptoe, A. (2012). Type D personality, self-efficacy, and medication adherence following an acute coronary syndrome. *Psychosomatic Medicine*, *74*(1), 100-106.

Mommersteeg, P. M., Kupper, N., & Denollet, J. (2010). Type D personality is associated with increased metabolic syndrome prevalence and an unhealthy lifestyle in a cross-sectional Dutch community sample. *BMC Public Health*, *10*, 714-2458-10-714.

Mommersteeg, P. M., Pelle, A. J., Ramakers, C., Szabó, B. M., Denollet, J., & Kupper, N. (2012). Type D personality and course of health status over 18 months in outpatients with heart failure: multiple mediating inflammatory biomarkers. *Brain, Behavior, And Immunity*, *26*(2), 301-310.

Montero, P., Rueda, B., & Bermúdez, J. (2012). Relación de la personalidad tipo D y el agotamiento vital con las emociones negativas y el ajuste psicológico a la enfermedad cardiaca. [Relationship of type D personality and vital exhaustion with negative emotions and psychological adjustment to heart disease]. *Revista de Psicopatología y Psicología Clínica*, *17*(2), 93-106.

O'Dell, K. R., Masters, K. S., Spielmans, G. I., & Maisto, S. A. (2011). Does type-D personality predict outcomes among patients with cardiovascular disease? A meta-analytic review. *Journal of Psychosomatic Research*, *71*(4), 199-206.

O'Gara, P. T., Kushner, F. G., Ascheim, D. D., Casey, D. E., Chung, M. K., de Lemos, J. A., ... & Granger, C. B. (2013). 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology*, *61*(4), e78-e140.

Panel NCEPNE (2002). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation*, *106*(25), 3143-3421.

Park, H., Yoon, C., Kang, S., Choi, D., Kim, H., Cho, M. C., et al. (2013). Early-and late-term clinical outcome and their predictors in patients with ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction. *International Journal of Cardiology*, *169*(4), 254-261.

Pedersen, S. S., & Denollet, J. (2003). Type D personality, cardiac events, and impaired quality of life: a review. *European Journal of Cardiovascular Prevention & Rehabilitation*, *10*(4), 241-248.

Pedersen, S. S., Holkamp, P. G., Caliskan, K., van Domburg, R. T., Erdman, R. A., & Balk, A. H. (2006). Type D personality is associated with impaired health-related quality of life 7 years following heart transplantation. *Journal of Psychosomatic Research*, *61*(6), 791-795.

Pedersen, S. S., Lemos, P. A., van Vooren, P. R., Liu, T. K., Daemen, J., Erdman, R. A., et al. (2004). Type D personality predicts death or myocardial infarction after bare metal stent or sirolimus-eluting stent implantation: A rapamycin-eluting stent evaluated at rotterdam cardiology hospital (RESEARCH) registry substudy. *Journal of the American College of Cardiology*, *44*(5), 997-1001.

Pedersen, S. S., & Denollet, J. (2006). Is type D personality here to stay? Emerging evidence across cardiovascular disease patient groups. *Current Cardiology Reviews*, *2*(3), 205-213.

Pedersen, S. S., Theuns, D. A., Muskens-Heemskerk, A., Erdman, R. A., & Jordaens, L. (2007). Type-D personality but not implantable cardioverter-defibrillator indication is associated with impaired health-related quality of life 3 months post-implantation. *Europace*, *9*(8), 675-680.

Pelle, A. J., Schiffer, A. A., Smith, O. R., Widdershoven, J. W., & Denollet, J. (2010). Inadequate consultation behavior modulates the relationship between type D personality and impaired health status in chronic heart failure. *International Journal of Cardiology*, *142*(1), 65-71.

Pelle, A. J., Pedersen, S. S., Schiffer, A. A., Szabo, B., Widdershoven, J. W., & Denollet, J. (2010). Psychological distress and mortality in systolic heart failure. *Circulation*, *3*(2), 261-267.

Pelle, A. J., van den Broek, K. C., Szabó, B., & Kupper, N. (2010). The relationship between Type D personality and chronic heart failure is not confounded by disease severity as assessed by BNP. *International Journal Of Cardiology*, *145*(1), 82-83.

Perk, J., De Backer, G., Gohlke, H., Graham, I., Reiner, Z., Verschuren, M., et al. (2012). European guidelines on cardiovascular disease prevention in clinical practice (version 2012). *European Heart Journal*, *33*, 1635-1701.

Petrova, D., Garcia-Retamero, R., & Catena, A. (2015). Lonely hearts don't get checked: On the role of social support in screening for cardiovascular risk. *Preventive Medicine*, *81*, 202-208.

Polman, R., Borkoles, E., & Nicholls, A. R. (2010). Type D personality, stress, and symptoms of burnout: The influence of avoidance coping and social support. *British Journal of Health Psychology*, 1*5*(3), 681-696.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879-891.

Rosengren, A., Hawken, S., Ôunpuu, S., Sliwa, K., Zubaid, M., Almahmeed, W. A., et al. (2004). Association of psychosocial risk factors with risk of acute myocardial infarction in 11 119 cases and 13 648 controls from 52 countries (the INTERHEART study): Case-control study. *The Lancet*, *364*(9438), 953-962.

Schiffer, A. A., Pedersen, S. S., Widdershoven, J. W., & Denollet, J. (2008). Type D personality and depressive symptoms are independent predictors of impaired health status in chronic heart failure. *European Journal of Heart Failure*, *10*(8), 802-810.

Shanmugasegaram, S., Flett, G. L., Madan, M., Oh, P., Marzolini, S., Reitav, J., ... & Sturman, E. D. (2014). Perfectionism, Type D personality, and illness-related coping styles in cardiac rehabilitation patients. *Journal of Health Psychology*, *19*(3), 417-426.

Schiffer, A. A., Denollet, J., Widdershoven, J. W., Hendriks, E. H., & Smith, O. R. (2007). Failure to consult for symptoms of heart failure in patients with a type-D personality. *Heart*, *93*(7), 814-818.

Schiffer, A. A., Pedersen, S. S., Widdershoven, J. W., Hendriks, E. H., Winter, J. B., & Denollet, J. (2005). The distressed (type D) personality is independently associated with impaired health status and increased depressive symptoms in chronic heart failure. *European Journal of Cardiovascular Prevention and Rehabilitation*, *12*(4), 341-346.

Sher, L. (2005). Type D personality: The heart, stress, and cortisol. *Quarterly Journal of Medicine*, *98*(5), 323-329.

Smith, S. C., Allen, J., Blair, S. N., Bonow, R. O., Brass, L. M., Fonarow, G. C., et al. (2006). AHA/ACC guidelines for secondary prevention for patients with coronary and other atherosclerotic vascular disease: 2006 update. *Journal of the American College of Cardiology*, *47*(10), 2130-2139.

Steg, G., James, S. K., Atar, D., Badano, L. P., Blomstrom Lundqvist, C., Borger, M. A., ... & Gershlick, A. H. (2013). Guía de práctica clínica de la ESC para el manejo del infarto agudo de miocardio en pacientes con elevación del segmento ST.[Clinical guide of the Spanish Society of Cardiology for managing acute myocardial infarction in patients with ST segment elevation]. *Revista Española de Cardiología*, *66*(01), 53-53.

Stuckey, M. I. (2013). *Associations between heart rate variability and metabolic syndrome risk factors.* Electronic Thesis and Dissertation Repository. Paper 1293.

Svansdottir, E., Karlsson, H. D., Gudnason, T., Olason, D. T., Thorgilsson, H., Sigtryggsdottir, U., et al. (2012). Validity of type D personality in

Iceland: Association with disease severity and risk markers in cardiac patients. *Journal of Behavioral Medicine*, *35*(2), 155-166.

Svansdottir, E., Denollet, J., Thorsson, B., Gudnason, T., Halldorsdottir, S., Gudnason, V., et al. (2013). Association of type D personality with unhealthy lifestyle, and estimated risk of coronary events in the general icelandic population. *European Journal of Preventive Cardiology*, *20*(2), 322-330.

Thygesen, K., Alpert, J. S., Jaffe, A. S., White, H. D., Simoons, M. L., Chaitman, B. R., et al. (2012). Third universal definition of myocardial infarction. *Journal of the American College of Cardiology*, *60*(16), 1581-1598.

van den Broek, Krista C, Nyklícek, I., van der Voort, Pepijn H, Alings, M., Meijer, A., & Denollet, J. (2009). Risk of ventricular arrhythmia after implantable defibrillator treatment in anxious type D patients. *Journal of the American College of Cardiology*, *54*(6), 531-537.

Volz, A., Schmid, J., Zwahlen, M., Kohls, S., Saner, H., & Barth, J. (2011). Predictors of readmission and health related quality of life in patients with chronic heart failure: A comparison of different psychosocial aspects. *Journal of Behavioral Medicine*, *34*(1), 13-22.

von Känel, R., Barth, J., Kohls, S., Saner, H., Znoj, H., Saner, G., et al. (2009). Heart rate recovery after exercise in chronic heart failure: Role of vital exhaustion and type D personality. *Journal of Cardiology*, *53*(2), 248-256.

Whitehead, D. L., Perkins-Porras, L., Strike, P. C., Magid, K., & Steptoe, A. (2007). Cortisol awakening response is elevated in acute coronary syndrome patients with type-D personality. *Journal of Psychosomatic Research*, *62*(4), 419-425.

Williams, L., O'Carroll, R. E., & O'Connor, R. C. (2009). Type D personality and cardiac output in response to stress. *Psychology and Health*, *24*(5), 489-500.

Williams, L., O’Connor, R. C., Grubb, N., & O’Carroll, R. (2011a). Type D personality predicts poor medication adherence in myocardial infarction patients. *Psychology & Health*, *26*(6), 703-712.

Williams, L., O'Connor, R. C., Grubb, N. R., & O'Carroll, R. E. (2011b). Type D personality and illness perceptions in myocardial infarction patients. *Journal of Psychosomatic Research*, *70*(2), 141-144.

Williams, L., O'Connor, R. C., Howard, S., Hughes, B. M., Johnston, D. W., Hay, J. L., et al. (2008). Type-D personality mechanisms of effect: The role of health-related behavior and social support. *Journal of Psychosomatic Research*, *64*(1), 63-69.

Williams, L., Abbott, C., & Kerr, R. (2015). Health behaviour mediates the relationship between type D personality and subjective health in the general population. *Journal of Health Psychology.* Epub ahead of print. doi: 10.1177/1359105315571977

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361-370.

CHAPTER 8:
Social support predicts screening for cardiovascular risk

The content of this chapter has been published as Petrova, D., Garcia-Retamero, R., & Catena, A. (2015). Lonely hearts don't get checked: On the role of social support in screening for cardiovascular risk. *Preventive Medicine*, *81*, 202-208. Supplementary data are available here: http://tinyurl.com/LonelyHeartsPaper

Chapter 8

Lonely hearts don't get checked:
On the role of social support in screening for cardiovascular risk

Regular cardiovascular risk screening can prevent cardiovascular disease through timely implementation of lifestyle changes or medication. However, few studies have investigated what factors promote regular screening for risk factors like hypertension and high blood cholesterol. The aim of this study was to investigate the relationship between social support and adherence to cardiovascular risk screening. We analyzed data from the Spanish National Health Survey—a cross-sectional representative survey conducted by the Spanish Ministry of Health in 2012 (N=21 007). Participants reported whether they had their blood pressure and cholesterol levels measured by a health professional in the previous 12 months. Social support (i.e., the perception that emotional and practical support was available when needed) was measured with a validated scale. Multiple logistic regressions were conducted adjusted for demographic and health-related factors. Compared to individuals who reported a lack of social support, individuals who perceived sufficient social support were on average twice more likely to report participation in blood pressure screening, OR=2.06, 95% CI [1.60, 2.66] and cholesterol screening, OR=2.85, 95% CI [1.99, 4.09]. These effects were uniform across different demographics and were replicated in a previous wave of the survey. Factors associated with worse screening adherence were low social class, being single or widowed, smoking, alcohol consumption, and no history of cardiovascular risk. Perceptions of social support are positively related to cardiovascular risk screening adherence. Future research should investigate what type of social support most effectively increases screening participation among high risk populations.

1. Introduction

Cardiovascular disease (CVD) is currently the world's number one killer. In 2012 it was responsible for 3 out of every 10 deaths (World Health Organization, 2014). Regular screening for cardiovascular risk can significantly reduce the burden of CVD by timely implementation of lifestyle changes or medication (World Health Organization & UNAIDS, 2007). However, screening uptake rates in Europe are suboptimal, especially among those at high risk and of lower socio-economic status (Filippidis et al., 2014; Galán et al., 2006; Kaplan & Keil, 1993; Rodin et al., 2012; Rodríguez-Artalejo et al., 2003). Although cardiovascular screening is often part of the annual medical visit in Europe, medical visits alone may be insufficient to ensure optimal screening coverage and equity, with healthcare professionals potentially contributing to screening inequalities (Rodríguez-Artalejo et al., 2003). In view of these results, research into what social, behavioral, or motivational factors can increase cardiovascular screening adherence is needed. However, to the best of our knowledge there is little empirical evidence available. One exception is a study conducted by Ashida and colleagues (2010) who showed that social encouragement is related to increased intentions to undergo blood pressure, cholesterol, and glucose screening in Mexican-American adults. These results suggest that social support from others can have a positive influence on cardiovascular screening adherence.

An increasing amount of research in social epidemiology shows that social support protects from CVD. Compared to people who have less social support, people who have more social support are at a smaller cardiovascular risk: they are less likely to develop CVD and less likely to die from CVD if they develop it (Barth et al., 2010). Theories postulate that one of the ways in which social support promotes health is through encouraging positive health-related behaviors (Cohen, 1988; Gallant, 2013; Kouvonen et al., 2012; Shiovitz-Ezra & Litwin, 2012; Uchino, 2009). This suggests that social support can facilitate adherence to regular screening, thereby reducing cardiovascular risk.

More evidence for the role of social support in preventive behavior comes from research on cancer screening adherence. Both structural and functional social support have been related to greater adherence to breast (Allen et al., 1999; Katapodi et al., 2002; Taylor et al., 1998) and colorectal cancer screening (Honda & Kagawa-Singer, 2006; Kinney et al., 2005). Structural support refers to the characteristics of one's social network (Allen et al., 1999; Keating et al., 2011; Kinney et al., 2005), while functional support refers to the perception that support is available when needed (Wills & Ainette, 2012). For example, people are more likely to participate in cancer screening if people from their social network participated before (Keating et al., 2011) or if the social network approves of screening (Allen et al., 1999; Honda & Kagawa-Singer, 2006). However, there are substantial differences between cancer screening and cardiovascular screening, and the unique role of social support in cardiovascular screening still needs to be established.

Social support can promote screening through provision of information about the importance or availability of screening (informational support), the opportunity to discuss openly about personal health concerns or doubts (emotional support), or practical assistance like driving someone to the health centre (instrumental support) (Fischer Aggarwal et al., 2008; Wills & Ainette, 2012). Screening for cardiovascular risk is recommended to start as early as age 20, and should be repeated at least every two years or more often depending on results (Greenland et al., 2010). In addition, research shows that regular medical visits alone cannot ensure screening (Rodríguez-Artalejo et al., 2003). This implies that an additional initiative from the side of the patient may be required for screening completion. Altogether the need for frequent and regular checks and the need for an initiative from the patient suggest that informational, emotional, and instrumental social support can have a strong positive influence on cardiovascular screening adherence. This suggests that social support will be positively related to cardiovascular screening adherence, such that people who lack social support are less likely to get screened, while people who have sufficient social support successfully adhere to screening guidelines.

Alternatively, social support may only be beneficial when screening is more invasive or the screening results are potentially more psychologically distressing. To illustrate, cancer screenings often involve invasive and unpleasant testing (e.g., colonoscopy and mammography). In addition, the presence of a positive result can indicate the presence of a severe diagnosis. No wonder that under such conditions, greater perceived support from friends or family has been related to more frequent cancer screening (Allen et al., 1999; Honda & Kagawa-Singer, 2006; Katapodi et al., 2002). Cardiovascular screening, on the other hand, involves relatively non-invasive techniques like blood pressure measurement and a simple blood test. In addition, a positive result only indicates the presence of an elevated risk factor as opposed to a severe diagnosis. Thus, beneficial effects of social support may be restricted to more invasive and psychologically distressing testing and may not exist for more harmless and simple tests like those that screen for cardiovascular risk.

Finally, the provision of social support can have beneficial effects on preventive health behavior only when it is perceived as constructive and positive. Social network members can fail to provide the needed support, can exert negative influence on health behavior, or provide assistance that is perceived as intrusive, causing negative effects on health (Rook, 2015; Taylor, 2011). To investigate the influence of constructive social support, in this research we used a measure of *perceived* social support, defined as the satisfaction with the functional and affective aspects of one’s social network (e.g., the perception that emotional and instrumental support is available when needed) (Broadhead et al., 1988).

We investigated if social support was related to cardiovascular screening adherence in probabilistic national samples (i.e., general population) in a South European country—Spain. We aimed to quantify the relationship between social support and cardiovascular risk screening after controlling for the effect of socio-demographic factors and other potential predictors of screening adherence.

2. Method

We obtained data from the adult National Health Survey (NHS) conducted by the Spanish Ministry of Health, Social Services and Equality, and the National Statistical Institute. The survey is a part of a periodic cross-sectional survey wave investigating health outcomes in Spanish citizens that is part of the European Health Survey project.[9]

Our primary data set was the NHS 2012, conducted between July 2011 and June 2012. The survey covered the whole territory of Spain. Multi-stage stratified random sampling was used to obtain a representative sample of the Spanish adult population (15 years or older) (Spanish Ministry of Health, Social Services and Equality & National Statistical Institute of Spain, 2012). To be able to provide reliable national estimates, 24 000 households distributed across 2 000 census sections were contacted. The census sections were initially grouped in 7 strata according to the size of the municipality to which they pertained and were selected with a probability proportional to this size. Twelve households were then randomly selected within each census section. If the selected household could not be included (e.g., it was empty or the residents refused to participate), the missing case was replaced with one with

---

[9] More information and technical details about the sampling, measures, and execution of the surveys are available on the website of the Ministry of Health: http://www.msssi.gob.es/estadEstudios/estadisticas/encuestaNacional.

similar characteristics from a replacement household sample. From each household one adult was randomly selected to participate in the survey. Data were collected via a computer-assisted personal interview by trained interviewers. Data collection was uniformly distributed across the 12-month period. The response rate was 90% and the final public data set included data from 21 007 adults.

We used the results of a previous wave of the survey, NHS 2007, conducted between June 2006 and June 2007, to cross-validate the results from NHS 2012. NHS 2007 employed similar methodology to NHS 2012, with the difference that it did not contain data on marital status, had a larger percentage of missing data, and the interview was not computer-assisted. In NHS 2007, 31 300 households distributed across 2 236 census sections were initially contacted. The response rate was 96% and the final public data set included data from 29 478 adults.

Both surveys included similar modules covering information about health and the use of health services. For this research, we obtained the measures described below. Unless otherwise specified, the two surveys used the same instruments and questions.

2.1. Measures

Social support. Social support was measured with the validated Spanish version of the Duke-UNC Functional Social Support Questionnaire (Bellon Saameno et al., 1996; Broadhead et al., 1988). The questionnaire consists of 11 items measuring functional and qualitative aspects of perceived social, emotional, and practical support (examples: "I get useful advice about important things in life", "I get invitations to go out and do things with other people", "I receive help when I am sick in bed"). Responses ranged from 1 (much less than I would like) to 5 (as much as I would like). The final score is a sum of all items, where a larger score indicates more social support (Cronbach's alpha NHS 2012=.91; NHS 2007=.96). Factor analysis using the principal components methods and a Varimax rotation resulted in the extraction of two components consistent with previous research (Bellon Saameno et al., 1996; Broadhead et al., 1988): affective/confident support (7 items) and functional/connectedness (4 items), explaining 66% of the variance in NHS 2012 and 63% in NHS 2007.

Cardiovascular screening. Participants indicated if they had their (1) blood pressure and (2) cholesterol level measured by a health professional in the past 12 months.

Control variables. Participants indicated if they had ever suffered from a chronic disease[10], diabetes, hypertension, high cholesterol, and heart disease (yes or no). Participants indicated their weight and height, from which we computed a body-mass index (BMI = weight in kg/(height in m)^2). In addition, participants indicated their smoking habits on a scale ranging (1: I smoke daily, 2: I smoke but not daily, 3: I currently do not smoke but have smoked before, and 4: I do not smoke and have never smoked habitually) (Spanish Ministry of Health, Social Services and Equality & National Statistical Institute of Spain, 2012). Participants reported their alcohol consumption for the past 12 months per week in grams[11]. Participants also rated their perceptions of health in the past 12 months on a scale from 1 (very bad) to 5

---

[10] NHS 2012: Participants were asked if they have ever suffered from a chronic disease. NHS 2007: This question was not included in this survey. Instead, participants indicated if they have suffered from each one of 29 chronic conditions. Hence, participants were coded as having suffered a chronic disease if they had indicated yes for any of these conditions.

[11] NHS 2012: Participants reported their alcohol consumption for the past 12 months (e.g., number of glasses of beer, wine, aperitifs, liquors, whisky, cider, etc., consumed per week). From these a measure of the mean consumption of pure alcohol per week in grams was calculated, using estimated values of pure alcohol in a glass for each type of drink. NHS 2007: This survey only included a measure of the overall frequency of alcohol consumption of various alcoholic drinks. In particular, participants were first asked if they have ever consumed alcohol. In case of affirmative answer, they were asked how often they consumed (1) wine, (2) beer with alcohol, (3) aperitifs with alcohol (e.g., vermouth), (4) cider, (5) mixed alcoholic drinks, brandy or liquor, and (6) whisky on scales from 0 (never) to 5 (daily). A composite score was created by computing the sum of these questions, whereby people who never consumed alcohol received a score of 0. Larger scores indicated more frequent consumption of alcoholic drinks.

(very good). Mental health was measured with the general Health GHQ-12 questionnaire (Golderberg & Williams, 1988), which has been validated in a Spanish population (López & Dresch, 2008). The instrument consists of 12 questions about the person's mental health in the past month compared to usually (e.g., if the person has experienced some symptoms or behavioral changes, "Have you had the feeling that you cannot overcome the difficulties you are facing?"). Each question is scored as 0 (not more than usually) or 1 (more than usually). The final score is a sum of all items, where a larger score indicates worse mental health.

Finally, participants' age, sex, marital status (married, single, widowed, separated, or divorced), and social class were recorded. Regarding social class, participants were classified into 6 categories (see Table 1S) according to the guidelines of the Spanish Society of Epidemiology (Domingo-Salvany et al., 2013).

2.2. Analysis plan.

We first analyzed the results from NHS 2012 and then sought to replicate the obtained results in NHS 2007. We conducted multiple logistic regressions with blood pressure screening and cholesterol screening as dependent variables. The main predictor of interest was social support. We also controlled for demographics and health history. Categorical predictors were modeled as dummy variables. To account for the sampling method and obtain unbiased parameter estimates representative of the population, we conducted logistic regressions for complex samples in SPSS. Sampling weights were provided with the survey data (Spanish Ministry of Health, Social Services and Equality & National Statistical Institute of Spain, 2012). We adopted an alpha level of .05 and used Bonferroni correction for multiple comparisons. The overall significance of each predictor was tested with the Wald F test. Odds ratios (OR) and 95% confidence intervals (CI) were calculated. In particular, to describe the relationship between social support and screening we calculated ORs for a 1-unit change in social support. To provide a more intuitive estimate of this effect, we also calculated ORs for a change over the entire range of the social support scale (i.e., 44 units), reflecting the difference in odds of screening between a lack of social support and sufficient social support.

3. Results

Descriptive statistics are included in Table 1 and Tables 1S and 2S of the online supplement. The results of the logistic regressions are displayed in Table 2.

3.1. Blood pressure screening

In NHS 2012 being older, belonging to middle social class (vs. lower), and being married (vs. single) was related to higher probability of screening. Having chronic disease, diabetes, previous history of high cholesterol, heart disease, or hypertension, a higher BMI and worse perceived health were also related to higher probability of screening. Ex-smokers were more likely to report screening than daily smokers and non-smokers. Finally, after accounting for the influence of all these factors, social support was significantly related to screening, OR=1.02, 95% CI [1.01, 1.02] (Figure 1). In particular, compared to people reporting a lack of social support, people who perceived sufficient social support were on average twice as likely to undergo blood pressure screening, OR=2.06, 95% CI [1.60, 2.66] (OR for 44 points of change, from lowest to highest social support).

The results for NHS 2007 were highly similar (see Table 2A). In this survey lower social class was consistently associated with lower odds of screening and daily and occasional smokers were less likely to report screening compared to ex-smokers and non-smokers.

Next, we checked if the influence of social support varied as a function of demographics by checking for significant interactions between social support and each of the demographic variables (sex, age, social class, marital status[12]) in both surveys. None of the tested interactions reached significance (p>.05), indicating that the relationship of social support to screening was uniform across different demographics.

3.2. Cholesterol screening

In NHS 2012 older age, being female, belonging to middle social class (vs. lower and higher), being married (as opposed to single or widowed) as well as being separated or divorced (vs. widowed) were related to higher probability of screening. In addition, having chronic disease, diabetes, previous history of high cholesterol or heart disease and worse perceived health were again related to higher probability of screening. Being an ex-smoker (vs. all other groups) was related to higher odds of screening. Finally, social support was significantly related to screening (Figure 1), OR=1.02, 95% CI [1.02, 1.03]. In particular, compared to people reporting a lack of social support, people who perceived sufficient social support were about three times as likely to undergo cholesterol screening, 44-point-change-OR=2.85, 95% CI [1.99, 4.09].

These results were replicated in NHS 2007. In addition, in this survey low social class and more alcohol consumption were associated with lower odds of screening while higher BMI was related to higher odds of screening. Again, we tested for interactions between social support and demographic variables. The only interaction that reached significance was that between social support and sex in NHS 2012, $F(1, 89)=4.98$, $p=.028$, indicating a stronger beneficial influence of social support on cholesterol screening adherence for women, 44-ponit-change OR=3.69, 95% CI [2.50, 5.43] than for men with 44-point-change-OR=2.14, 95% CI [1.31, 3.52].

---

[12] The interaction terms were computed with social support as a continuous variable and sex, social class, and marital status as categorical variables. Interactions with age were tested with age both as a continuous and dichotomized variable (cut-off 50).

Table 1. Screening and health history of participants (Spain, 2007, 2012).

	NHS 2012						NHS 2007					
	Yes		***No***		***Missing***		***Yes***		***No***		***Missing***	
	N	%	N	%	N	%	N	%	N	%	N	%
Blood pressure	15093	71.8	5875	28.0	39	.2	20874	70.7	6862	23.3	1769	6.0
Cholesterol screening*	13922	66.3	6943	33.1	142	.7	18154	61.6	9405	31.9	1919	6.5
Chronic disease	9906	47.2	11083	52.8	18	.1	23621	80.1	5699	19.3	158	.5
Diabetes	1870	8.9	19117	91.0	20	.1	2178	7.4	27213	92.3	87	.3
Hypertension history	5513	26.2	15452	73.6	42	.2	7575	25.7	21820	74.0	83	.3
High cholesterol history	4634	22.1	16301	77.6	72	.3	5564	18.9	23827	80.8	87	.3
Heart disease history	1915	9.1	19072	90.8	20	.1	2712	9.2	26673	90.5	93	.3

Note: *In the past 12 months

Table 2. Logistic regression results for participation in screening for cardiovascular risk in the previous 12 months (Spain, 2007, 2012).

Note: N=number of participants after accounting for missing values. Bonferroni corrected p value; OR=odds ratio; LLCI=95% lower level confidence interval; ULCI=95% upper level confidence interval. Significant ORs are marked in bold. The dependent variables were coded as 0=no screening, 1=screening. ORs for continuous predictors indicate the odds ratio for one unit change in the predictor, where OR>1 indicates that a higher score on the predictor variable was associated with higher odds of screening. For categorical predictors, OR<1 indicates that the reference category of the predictor was associated with lower odds of screening (e.g., no history of heart disease was associated with lower odds of screening compared to having history). Marital status was not recorded in NHS 2007 and could not be included in the models. Social class: 1 (highest; e.g., managers of big establishments), 2 (e.g., managers of small establishments), 3 (e.g., intermediate occupations), 4 (e.g., qualified technical workers), 5 (e.g., semi-qualified workers), 6 (lowest; e.g., unqualified workers).

A) BLOOD PRESSURE SCREENING

	NHS 2012 (N=14 865)					NHS 2007 (N=20 679)					
Correct classification	66.60%					73.00%					
Predictor	Wald F	p	OR	LLCI	ULCI	Wald F	p	OR	LLCI	ULCI	OR reference
Sex	0.84	0.361	0.95	0.86	1.06	1.01	0.318	0.94	0.84	1.06	male vs. female
Age	56.92	0.000	**1.01**	1.01	1.02	83.74	0.000	**1.02**	1.01	1.02	-
Social class	2.54	0.034	1.14	0.91	1.43	8.30	0.000	1.13	0.91	1.39	1 vs. 2
			1.03	0.89	1.19			1.17	0.98	1.40	1 vs. 3
			0.90	0.75	1.09			1.04	0.84	1.29	2 vs. 3
			0.94	0.78	1.13			**1.31**	1.06	1.62	1 vs. 4
			0.82	0.65	1.04			1.16	0.96	1.41	2 vs. 4
			0.91	0.76	1.10			1.12	0.95	1.31	3 vs. 4
			1.13	0.97	1.31			**1.45**	1.21	1.75	1 vs. 5
			0.99	0.80	1.22			**1.29**	1.07	1.57	2 vs. 5
			1.10	0.94	1.28			**1.24**	1.08	1.43	3 vs. 5
			1.20	1.04	1.38			1.11	0.96	1.29	4 vs. 5
			1.19	1.00	1.43			**1.65**	1.38	1.99	1 vs. 6 (low)
			1.05	0.83	1.32			**1.47**	1.24	1.74	2 vs. 6 (low)

Table 2 continued.

			1.16	0.96	1.40			**1.41**	1.19	1.68	3 vs. 6 (low)
			1.27	1.07	1.51			**1.26**	1.07	1.49	4 vs. 6 (low)
			1.06	0.90	1.24			1.14	0.99	1.30	5 vs. 6 (low)
Marital status	2.14	0.595	**0.86**	0.77	0.97	–	–				single vs. married
			0.88	0.68	1.15						widowed vs. married
			1.00	0.67	1.50						separated vs. married
			1.04	0.82	1.32						divorced vs. married
			1.02	0.76	1.37						widowed vs. single
			1.15	0.75	1.77						separated vs. single
			1.20	0.93	1.55						divorced vs. single
			1.13	0.67	1.91						separated vs. widowed
			1.18	0.89	1.55						divorced vs. widowed
			1.04	0.67	1.62						divorced vs. separated
Chronic disease	49.19	0.000	**0.71**	0.65	0.79	5.73	0.019	**0.87**	0.78	0.98	no vs. yes
Diabetes	7.99	0.006	**0.62**	0.44	0.87	22.16	0.000	**0.42**	0.29	0.60	no vs. yes
History of hypertension	26.23	0.000	**0.49**	0.37	0.65	223.64	0.000	**0.33**	0.28	0.38	no vs. yes
History of high cholesterol	41.01	0.000	**0.57**	0.48	0.68	36.88	0.000	**0.68**	0.60	0.77	no vs. yes
History of heart disease	12.78	0.001	**0.60**	0.45	0.80	17.21	0.000	**0.55**	0.41	0.73	no vs. yes
BMI	7.33	0.008	**1.02**	1.01	1.03	21.67	0.000	**1.03**	1.02	1.04	-
Smoking	3.75	0.053	1.26	0.99	1.61	4.71	0.089	0.93	0.74	1.16	occasionally vs. daily
			1.22	1.08	1.37			**1.26**	1.10	1.44	ex-smoker vs. daily
			1.05	0.94	1.18			**1.16**	1.02	1.32	non-smoker vs. daily
			0.97	0.78	1.20			**1.36**	1.08	1.71	ex-smoker vs. occasionally
			0.83	0.64	1.09			**1.25**	1.00	1.56	non-smoker vs. occasionally
			0.86	0.76	0.97			0.92	0.81	1.05	non-smoker vs. ex-smoker
Alcohol	1.78	0.185	1.00	0.99	1.00	0.13	0.718	1.00	0.99	1.01	-
Perceived health	29.05	0.000	**0.77**	0.70	0.85	35.87	0.000	**0.77**	0.70	0.84	-
Mental health	0.28	0.601	1.01	0.98	1.03	0.01	0.924	1.00	0.98	1.02	-
Social support	31.61	0.000	**1.02**	1.01	1.02	13.01	0.001	**1.01**	1.01	1.02	-

Table 2 continued.

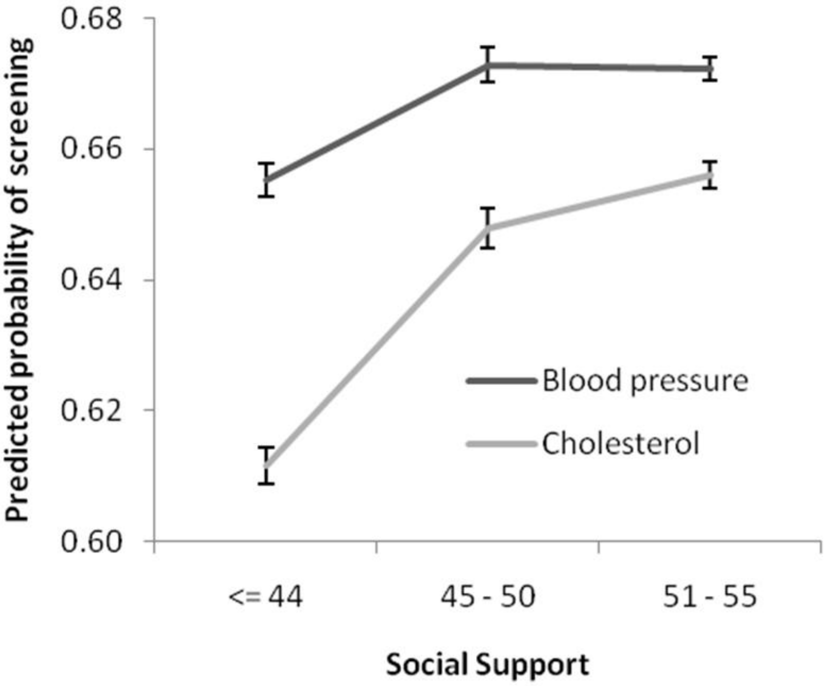
B) CHOLESTEROL SCREENING

	NHS 2012 (N=14 805)					NHS 2007 (N=20 629)					
Correct classification			0.66					0.66			
Predictor	*Wald F*	*p*	*OR*	*LLCI*	*ULCI*	*Wald F*	*p*	*OR*	*LLCI*	*ULCI*	*OR reference*
Sex	4.98	0.028	**0.89**	0.80	0.99	1.09	0.300	0.93	0.81	1.07	male vs. female
Age	76.28	0.000	**1.01**	1.01	1.02	56.54	0.000	**1.01**	1.01	1.02	-
Social class	5.22	0.001	1.10	0.90	1.35	7.71	0.000	0.98	0.79	1.20	1 vs. 2
			1.05	0.89	1.24			1.08	0.91	1.28	1 vs. 3
			0.96	0.80	1.15			1.10	0.92	1.32	2 vs. 3
			0.82	0.69	0.99			1.08	0.91	1.27	1 vs. 4
			0.75	0.63	0.89			1.10	0.90	1.36	2 vs. 4
			0.78	0.66	0.93			1.00	0.86	1.16	3 vs. 4
			1.11	0.96	1.29			**1.28**	1.09	1.51	1 vs. 5
			1.01	0.88	1.17			**1.31**	1.11	1.56	2 vs. 5
			1.06	0.93	1.20			**1.19**	1.04	1.36	3 vs. 5
			1.35	1.19	1.53			**1.19**	1.02	1.38	4 vs. 5
			1.14	0.95	1.36			**1.52**	1.27	1.82	1 vs. 6 (low)
			1.04	0.86	1.24			**1.56**	1.31	1.86	2 vs. 6 (low)
			1.08	0.91	1.28			**1.41**	1.21	1.64	3 vs. 6 (low)
			1.38	1.17	1.63			**1.41**	1.20	1.66	4 vs. 6 (low)
			1.02	0.88	1.20			**1.18**	1.04	1.35	5 vs. 6 (low)
Marital status	6.14	0.093	**0.81**	0.71	0.93	–	–				single vs. married
			0.69	0.53	0.90						widowed vs. married
			1.09	0.77	1.54						separated vs. married
			0.98	0.77	1.25						divorced vs. married
			0.85	0.61	1.18						widowed vs. single

Table 2 continued.

			1.34	0.92	1.97						separated vs. single
			1.21	0.91	1.61						divorced vs. single
			1.58	1.07	2.33						separated vs. widowed
			1.42	1.05	1.93						divorced vs. widowed
			0.90	0.65	1.26						divorced vs. separated
Chronic disease	32.51	0.000	**0.71**	0.63	0.80	7.17	0.009	**0.84**	0.73	0.96	no vs. yes
Diabetes	36.45	0.000	**0.50**	0.40	0.63	74.80	0.000	**0.44**	0.36	0.53	no vs. yes
History of hypertension	0.13	0.717	1.05	0.80	1.39	23.00	0.000	**0.73**	0.64	0.83	no vs. yes
History of high cholesterol	236.81	0.000	**0.28**	0.24	0.33	381.82	0.000	**0.27**	0.24	0.31	no vs. yes
History of heart disease	4.40	0.039	**0.81**	0.67	0.99	7.96	0.006	**0.78**	0.65	0.93	no vs. yes
BMI	6.70	0.011	1.02	1.00	1.03	40.75	0.000	**1.03**	1.02	1.04	-
Smoking	13.72	0.000	1.16	0.92	1.46	9.22	0.017	0.98	0.77	1.24	occasionally vs. daily
			1.48	1.30	1.68			**1.37**	1.21	1.55	ex-smoker vs. daily
			1.08	0.96	1.22			1.14	1.00	1.31	non-smoker vs. daily
			1.27	1.01	1.61			**1.40**	1.10	1.77	ex-smoker vs. occasionally
			0.93	0.73	1.18			1.17	0.94	1.46	non-smoker vs. occasionally
			0.73	0.65	0.83			**0.84**	0.74	0.95	non-smoker vs. ex-smoker
Alcohol	11.98	0.001	0.99	0.98	1.00	10.97	0.001	**0.98**	0.98	0.99	-
Perceived health	20.27	0.000	**0.81**	0.74	0.89	32.06	0.000	**0.86**	0.81	0.90	-
Mental health	0.31	0.581	0.99	0.97	1.02	1.31	0.255	1.01	0.99	1.04	-
Social support	33.15	0.000	**1.02**	1.02	1.03	15.50	0.000	**1.01**	1.01	1.02	-

Figure 1. Predicted probability of blood pressure and cholesterol screening obtained from the regression models in NHS 2012 as a function of social support. Error bars are ±1 standard error of the mean (Spain, 2012).



4. Discussion

We investigated the relationship between perceived social support and self-reported screening for cardiovascular risk in a survey conducted in probabilistic national samples in Spain. Across two types of cardiovascular screening, more perceptions of social support were related to higher odds of screening. Importantly, social support was a significant predictor of screening after accounting for relevant socio-demographic and health factors, and its influence was uniformly positive across different demographics. Our research demonstrates that social support can be beneficial even when the screening procedure is relatively non-distressing, non-invasive, inexpensive, and easily available. These results are consistent with previous research showing that the availability of supportive resources is related to greater screening adherence (Allen et al., 1999; Honda & Kagawa-Singer, 2006; Katapodi et al., 2002). An optimistic finding is that social support can be especially beneficial for women's adherence to cholesterol screening. In HNS 2012 women were more likely to report screening than men. However, previous research in the Basque country in Spain showed that women were less likely to have their cholesterol measured (4). To the extent that such inequalities still exist in some regions, they could be potentially reduced by social support interventions.

Consistent with previous research, individuals of lower socio-economic class were less likely to have been screened (Rodin et al., 2012; Rodríguez-Artalejo et al., 2003). Interestingly, in NHS 2012 individuals belonging to the middle class (i.e., qualified technical workers) were most likely to attend to cholesterol screening. We could speculate that this finding has to do with the economic crisis in Spain, which had its peak around the time this survey was conducted and had an especially strong effect on Spain's middle class. People from this class may have been more likely to experience stress and health problems and hence more likely to visit their doctor and get screened. It is possible that people who experience symptoms use the health system more often, thereby increasing the likelihood of screening for cardiovascular risk. Consistent with this hypothesis, individuals who perceived their health to be worse, had a larger BMI, and had previous history of cardiovascular risk were more likely to have been screened. More frequent screening of those at higher risk is an optimistic finding in line with international expert guidelines (Greenland et al., 2010).

However, another set of risk factors showed a worrisome trend. Consistent with previous studies of cardiovascular (Ashida et al., 2010; Filippidis et al., 2014) and cancer screening (Byrne et al., 2010; Fredman et al., 1999; MacLaughlan et al., 2011; Selvin & Brett, 2003), smoking and alcohol consumption were related to lower odds of screening. These results are especially worrisome because smoking and alcohol consumption drastically increase the risk of CVD and cancer (World Health Organization & UNAIDS, 2007). Prominent health behavior models like the Theory of Planned Behavior (Ajzen, 1991) or the Health Belief Model (Janz & Becker, 1984) suggest that these individuals may engage in risky behavior and forego preventive services because they do not feel at risk of CVD (Byrne et al., 2010). Indeed, both lower perceived risk and less attribution of CVD to lifestyle factors are related to fewer intentions to screen for cardiovascular risk (Ashida et al., 2010). Although daily smokers are at greater cardiovascular risk compared occasional smokers (Luoto et al., 2000), there were no significant differences in screening attendance between these groups. In contrast, ex-smokers were consistently more likely to report screening. This could be related to increased health conscientiousness in these individuals, which could have resulted in quitting smoking and regular attendance to screening. Alternatively, some of these individuals may have quit smoking after unfavorable results from screening. More research is needed into why people belonging to vulnerable socio-demographic and risk groups fail to adhere to screening recommendations and what types of social support can most effectively increase their participation.

Our study measured social support defined as the perception that emotional and instrumental support is available (Broadhead et al., 1988). This perception can be a result of a number of factors both external and internal to the individual, such as the size of one’s supportive network, having a close confidant, or a personal disposition to easily create relationships with others. In this vein, structural social support (contact with people in one’s social network) may provide the basis for functional social support to occur (these people might provide help and advice) (Hwang et al., 2014). For example, the presence of people who have undergone screening in one’s social network may create the impression that screening is a normative, desirable behavior (Allen et al., 1999; Honda & Kagawa-Singer, 2006). Other examples of more direct social influence and support are modifying beliefs concerning early detection and increasing health conscientiousness; decreasing difficulties related to cost, scheduling, and transportation; or discussing sensitive health issues with a friend (Allen et al., 1999; Honda & Kagawa-Singer, 2006; Seow et al., 2000; Taylor et al., 1998). Of particular interest for future research are specific mechanisms and support behaviors that could be addressed in interventions and campaigns (e.g., a campaign targeting the significant others of people at high risk to encourage them to do regular screening at home). Finally, different types of social support may be effective for different vulnerable groups (e.g., instrumental support for the elderly or economically disadvantaged individuals vs. emotional and informational support for smokers).

Like every research this study has several limitations. Previous studies have shown that uninsured individuals are more likely to forego screening (Byrne et al., 2010; Sabatino et al., 2008). While our results are likely to generalize to the Spanish population where virtually everyone is insured, it is not clear to what extent they would apply to citizens from countries where many people are uninsured (Schoen et al., 2005). While we controlled for a number of variables in our regression models, it is possible that there are other factors explaining the relationship between social support and cardiovascular screening. For example, health conscientiousness could be a potential mediator, such that receiving social support increases health conscientiousness and thus screening attendance. Alternatively, health conscientiousness could be a confounding factor that is not on the causal pathway. Social support was measured after screening attendance, so it is possible that screening attendance influenced perceptions of social support. We consider it less likely that one single activity in the past year strongly influenced global perceptions of social support. However, in reality social support and preventive health behavior may be bi-directionally related. Unfortunately, to the best of our knowledge, currently there is no longitudinal nationally representative data available that would allow testing this hypothesis. Future research should also record actual screening attendance, as opposed to trusting participants' self reports which could be susceptible to a number of biases. Finally, greater precision could be achieved by modeling the influence of education and income separately.

Experts have recommended regular screening for cardiovascular risk because it can reduce cardiovascular events without incurring any substantial harm to the individuals who participate (Sheridan et al., 2003). Consistent with theories about the effects of social support on health outcomes (Cohen, 1988; Gallant, 2013; Kouvonen et al., 2012; Shiovitz-Ezra & Litwin, 2012; Uchino, 2009), our results show that "lonely hearts don't get checked": People who lack social support and are thus at higher risk of cardiovascular disease (Barth et al., 2010) are also significantly less likely to use preventive services (Cohen, 1988; Gallant, 2013; Kouvonen et al., 2012; Shiovitz-Ezra & Litwin, 2012; Uchino, 2009). In a more optimistic perspective, these results suggest that provision of social support or increasing perceptions of social support can promote regular cardiovascular screening and thus decrease the risk of cardiovascular events. An essential step in future research is to investigate what specific supportive behaviors most effectively increase screening participation.

5. References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Allen, J. D., Sorensen, G., Stoddard, A. M., Peterson, K. E., & Colditz, G. (1999). The relationship between social network characteristics and breast cancer screening practices among employed women. *Annals of Behavioral Medicine, 21*(3), 193-200.

Ashida, S., Wilkinson, A. V., & Koehly, L. M. (2010). Motivation for health screening: Evaluation of social influence among mexican-american adults. *American Journal of Preventive Medicine, 38*(4), 396-402.

Barth, J., Schneider, S., & von Kanel, R. (2010). Lack of social support in the etiology and the prognosis of coronary heart disease: A systematic review and meta-analysis. *Psychosomatic Medicine, 72*(3), 229-238.

Bellon Saameno, J. A., Delgado Sanchez, A., Luna del Castillo, J. D., & Lardelli Claret, P. (1996). Validity and reliability of the duke-UNC-11 questionnaire of functional social support. [Validez y fiabilidad del cuestionario de apoyo social functional Duke-UNC-11] *Atencion Primaria, 18*(4), 153-6, 158-63.

Broadhead, W., Gehlbach, S. H., De Gruy, F. V., & Kaplan, B. H. (1988). The Duke-UNC functional social support questionnaire: Measurement of social support in family medicine patients. *Medical Care, 26*(7), 709-723.

Byrne, M. M., Davila, E. P., Zhao, W., Parker, D., Hooper, M. W., Caban-Martinez, A., . . . Lee, D. J. (2010). Cancer screening behaviors among smokers and non-smokers. *Cancer Epidemiology, 34*(5), 611-617.

Cohen, S. (1988). Psychosocial models of the role of social support in the etiology of physical disease. *Health Psychology, 7*(3), 269-297.

Domingo-Salvany, A., Bacigalupe, A., Carrasco, J. M., Espelt, A., Ferrando, J., & Borrell, C. (2013). Propuestas de clase social neoweberiana y neomarxista a partir de la clasificación nacional de ocupaciones 2011. *Gaceta Sanitaria, 27*(3), 263-272.

Filippidis, F. T., Gerovasili, V., & Majeed, A. (2014). Association between cardiovascular risk factors and measurements of blood pressure and cholesterol in 27 European countries in 2009. *Preventive Medicine, 67*, 71-74.

Fischer Aggarwal, B. A., Liao, M., & Mosca, L. (2008). Physical activity as a potential mechanism through which social support may reduce cardiovascular disease risk. *The Journal of Cardiovascular Nursing, 23*(2), 90-96.

Fredman, L., Sexton, M., Cui, Y., Althuis, M., Wehren, L., Hornbeck, P., & Kanarek, N. (1999). Cigarette smoking, alcohol consumption, and screening mammography among women ages 50 and older. *Preventive Medicine, 28*(4), 407-417.

Galán, I., Rodríguez-Artalejo, F., Díez-Gañán, L., Tobías, A., Zorrilla, B., & Gandarillas, A. (2006). Clustering of behavioural risk factors and compliance with clinical preventive recommendations in Spain. *Preventive Medicine, 42*(5), 343-347.

Gallant, M. P. (2013). Social networks, social support, and health-related behavior. In L. R. Martin, & M. R. DiMatteo (Eds.), *The Oxford Handbook of Health communication, Behavior Change, and Treatment adherence.* (pp. 305-322). New York: Oxford University Press.

Golderberg, D., & Williams, P. (1988). A user's guide to the general health questionnaire. *Windsor, UK: NFER-Nelson,*

Greenland, P., Alpert, J. S., Beller, G. A., Benjamin, E. J., Budoff, M. J., Fayad, Z. A., . . . Kushner, F. G. (2010). 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive Summary. A report of the American college of cardiology. *Journal of the American College of Cardiology, 56*(25), 2182-2199.

Honda, K., & Kagawa-Singer, M. (2006). Cognitive mediators linking social support networks to colorectal cancer screening adherence. *Journal of Behavioral Medicine, 29*(5), 449-460.

Hwang, K. O., Etchegaray, J. M., Sciamanna, C. N., Bernstam, E. V., & Thomas, E. J. (2014). Structural social support predicts functional social support in an online weight loss programme. *Health Expectations, 17*(3), 345-352.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education & Behavior, 11*(1), 1-47.

Kaplan, G. A., & Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: A review of the literature. *Circulation, 88*(4 Pt 1), 1973-1998.

Katapodi, M. C., Facione, N. C., Miaskowski, C., Dodd, M. J., & Waters, C. (2002). The influence of social support on breast cancer screening in a multicultural community sample. *Oncology Nursing Forum, 29*(5), 845-852.

Keating, N. L., O'Malley, A. J., Murabito, J. M., Smith, K. P., & Christakis, N. A. (2011). Minimal social network effects evident in cancer screening behavior. *Cancer, 117*(13), 3045-3052.

Kinney, A. Y., Bloor, L. E., Martin, C., & Sandler, R. S. (2005). Social ties and colorectal cancer screening among blacks and whites in North Carolina. *Cancer Epidemiology, Biomarkers & Prevention, 14*(1), 182-189.

Kouvonen, A., De Vogli, R., Stafford, M., Shipley, M. J., Marmot, M. G., Cox, T., . . . Kivimäki, M. (2012). Social support and the likelihood of maintaining and improving levels of physical activity: The Whitehall II study. *The European Journal of Public Health, 22*(4), 514-518.

López, M. P. S., & Dresch, V. (2008). The 12-item general health questionnaire (GHQ-12): Reliability, external validity and factor structure in the spanish population. *Psicothema, 20*(4), 839-843.

Luoto, R., Uutela, A., & Puska, P. (2000). Occasional smoking increases total and cardiovascular mortality among men. *Nicotine & Tobacco Research, 2*(2), 133-139.

MacLaughlan, S. D., Lachance, J. A., & Gjelsvik, A. (2011). Correlation between smoking status and cervical cancer screening: A cross-sectional study. *Journal of Lower Genital Tract Disease, 15*(2), 114-119.

Rodin, D., Stirbu, I., Ekholm, O., Dzurova, D., Costa, G., Mackenbach, J. P., & Kunst, A. E. (2012). Educational inequalities in blood pressure and cholesterol screening in nine european countries. *Journal of Epidemiology and Community Health, 66*(11), 1050-1055.

Rodríguez-Artalejo, F., Díez-Gañán, L., Artiñano, A. B., & Banegas Banegas, J. R. (2003). Effectiveness and equity of serum cholesterol and blood pressure testing: A population-based study in spain. *Preventive Medicine, 37*(2), 82-91.

Rook, K. S. (2015). Social networks in later life weighing positive and negative effects on health and well-being. *Current Directions in Psychological Science, 24*(1), 45-51.

Sabatino, S. A., Coates, R. J., Uhler, R. J., Breen, N., Tangka, F., & Shaw, K. M. (2008). Disparities in mammography use among US women aged 40-64 years, by race, ethnicity, income, and health insurance status, 1993 and 2005. *Medical Care, 46*(7), 692-700.

Schoen, C., Doty, M. M., Collins, S. R., & Holmgren, A. L. (2005). Insured but not protected: How many adults are underinsured? *Health Affairs, 24*(4), W5289-W5302.

Selvin, E., & Brett, K. M. (2003). Breast and cervical cancer screening: Sociodemographic predictors among white, black, and Hispanic women. *American Journal of Public Health, 93*(4), 618-623.

Seow, A., Huang, J., & Straughan, P. T. (2000). Effects of social support, regular physician and health-related attitudes on cervical cancer screening in an asian population. *Cancer Causes & Control, 11*(3), 223-230.

Sheridan, S., Pignone, M., & Donahue, K. (2003). Screening for high blood pressure: A review of the evidence for the US preventive services task force. *American Journal of Preventive Medicine, 25*(2), 151-158.

Shiovitz-Ezra, S., & Litwin, H. (2012). Social network type and health-related behaviors: Evidence from an American national survey. *Social Science & Medicine, 75*(5), 901-904.

Spanish Ministry of Health, Social Services and Equality, & National Statistical Institute of Spain. (2012). *National health survey 2011-2012: Methodology [encuesta nacional de salud 2011-2012: Metodología].*

Taylor, S. E. (2011). Social support: A review. In H. Friedman (Ed.), *Oxford Handbook of health psychology* (pp. 189-214). New York: Oxford University Press.

Taylor, V. M., Thompson, B., Montano, D. E., Mahloch, J., Johnson, K., & Li, S. (1998). Mammography use among women attending an inner-city clinic. *Journal of Cancer Education, 13*(2), 96-101.

Uchino, B. N. (2009). Understanding the links between social support and physical health: A life-span perspective with emphasis on the separability of perceived and received support. *Perspectives on Psychological Science, 4*(3), 236-255.

Wills, T. A., & Ainette, M. G. (2012). Social networks and social support. In A. Baum, A. Revenson & J. Singer (Eds.), *Handbook of health psychology* (pp. 465-492). New York: Psychology Press.

World Health Organization. (2014). *The top 10 causes of death.* Fact sheet N°310.

World Health Organization, & UNAIDS. (2007). *Prevention of cardiovascular disease*. Geneva: World Health Organization.

CHAPTER 9:
Effective evidence-based programs for prevention

The content of this chapter has been published as Petrova, D., & Garcia-Retamero, R. (2015). Effective evidence-based programs for preventing sexually-transmitted diseases: A meta-analysis. *Current HIV Research*, *13*(5), 432-438.

Chapter 9

Effective Evidence-Based Programs for Preventing Sexually-Transmitted Diseases: A Meta-Analysis

Educational programs for preventing sexually transmitted diseases (STDs) have often been implemented in different settings and populations. Mathematica Policy Research and Child Trends conducted a systematic review of 289 evidence-based interventions aiming to reduce STDs and sexual risk behavior in adolescents in the United States. These interventions were published between 1989 and 2012. We conducted a meta-analysis of the interventions that assessed incidence of STDs at follow up, and we identified key characteristics of successful interventions. Results showed that on average interventions reduced incidence roughly from 7 to 6 out of 100 people (17% relative risk reduction (RRR)). Interventions focused on abstinence had no effect, while comprehensive education programs aiming to improve skills and promote safe sexual practices reduced risk by 4 percentage points (23% RRR). In particular, interventions teaching condom use skills or communication and negotiation skills reduced incidence of STDs by 3 to 4 percentage points (30% RRR). Finally, interventions decreasing frequency of intercourse or number of sexual partners and interventions increasing condom use also reduced incidence of STDs by 5 to 7 percentage points (28-36% RRR). Overall properly designed interventions with the above-mentioned characteristics can achieve a 30% reduction of STD incidence. Implications for designing successful interventions to prevent STDs in adolescents are discussed.

1. Introduction

Sexually transmitted diseases (STDs) continue to be a burden in the United States, where about 20 million new infections occur every year (Center for Disease Control and Prevention (CDC), 2013). Teenagers and young adults are at highest risk of contracting STDs. In fact, 50% of STDs in the United States occur in people aged 15 to 24 (CDC, 2013). Despite the high prevalence of STDs, most teenagers and young adults do not believe they are at risk and often engage in risky sexual behaviors (Wildsmith, Schelar, Peterson, & Manlova, 2010; Garcia-Retameto & Cokely, 2011, 2015). Prevention of STDs can reduce the increasing rates and the corresponding health-care costs. As STDs begin to be acquired soon after sexual initiation, early education and prevention among adolescents is strongly recommended (Forhan et al., 2009).

Numerous interventions have been developed to reduce sexual risk taking among adolescents in the United States. These interventions range from programs promoting abstinence before marriage (e.g., Trenhol et al., 2007) to comprehensive sex education programs aiming to improve skills and promote safe sexual practices (e.g., DiClemente et al., 2004; Trenholm et al., 2007). Several authors (see DiCenso, Guyatt, Willan, & Griffith, 2002; Kirby & Laris, 2009; Kirby, Laris, & Rolleri, 2007; Picot et al., 2012; Sales, Milhausen, & DiClemente, 2006) conducted extensive systematic reviews of the efficacy of these interventions. These reviews focused on different outcome behaviors, criteria for study selection, and methods of data analysis. Frequency of sexual activity, use of contraception, and number of sexual partners were the most common outcome behaviors measured. Some reviews sampled studies according to the setting (e.g., interventions conducted in schools only (Silva, 2002) or goal (e.g., interventions specifically aiming to reduce unintended pregnancy DiCenso et al., 2002). Common methods of data analysis were systematic reviews (e.g., Picot et al., 2012) and meta-analyses (e.g., Johnson, Carey, Marsh, Levin, & Scott-Sheldon, 2003; Johnson, Scott-Sheldon, Huedo-Medina, & Carey, 2011). Despite differences, these reviews converge to suggest that interventions tend to have a modest effect on risky sexual behaviors (Kirby et al., 2007; Picot et al., 2012; Johnson et al., 2011). To illustrate, Johnson et al. (2011) conducted an extensive meta-analysis of 98 interventions aiming to reduce the incidence of the Human Immunodeficiency Virus (HIV) and related risky sexual behaviors. Results showed that overall interventions effectively increased condom use, reduced or delayed sex, and improved communication skills. There were also important moderators of these effects. For instance, interventions reduced the frequency of sexual intercourse if they promoted condom use and included motivational training rather than abstinence education. Johnson et al.'s review (2011) also concluded that these interventions can reduce incidence of STDs. However, the authors did not examine which variables explained the effect of these programs. In addition, they only selected studies that emphasized HIV content and excluded several abstinence or pregnancy prevention programs that might have affected incidence of STDs.

To the best of our knowledge, there is no recent systematic review examining the effectiveness of a broader spectrum of interventions to reduce STD incidence in US adolescents. The number of interventions conducted in the recent years is growing, which would allow to go beyond a narrative review of empirical evidence. Interventions with small samples are relatively frequent and often lack the power to detect effects (e.g., Roye, Silverman, & Krauss, 2007), rendering comparisons difficult or unfeasible. A meta-analytic review can include these studies and account for their power to detect a difference. In addition, most previous reviews focused on the impact of programs on knowledge, behavioral intentions, and risky behavior. However, none of these reviews investigated whether interventions translate into lower STD incidence and what are the factors or moderators contributing to this effect. In the current meta-analyses, we assessed the effectiveness of interventions to reduce incidence of STDs among US adolescents. We reviewed methodologically-sound programs aiming to reduce sexual risk behaviors and subjected them to a meta-analysis. We also examined several factors that could moderate the effectiveness of these programs. In particular, we examined the moderating role of (1) the characteristics of the intervention, including knowledge and skill content, and (2) changes in relevant sexual behaviors after the intervention (e.g., condom use and number of partners).

2. Method

2.1. Initial Study Sample

Mathematica Policy Research and its partner Child Trends were hired by the US Department of Health and Human Services to (1) conduct a systematic review of interventions designed to reduce sexual risk-taking among US adolescents and to (2) identify successful interventions (Mathematica Policy Research & Child Trends, 2012). The programs included in the last update of this review (i.e., in 2012; see also www.hhs.gov) comprised our initial study sample. Below we describe how studies were selected to enter Mathematica's review and how they were assessed for methodological quality. Further details can be obtained from the protocol of the review (see Mathematica Policy Research & Child Trends, 2012).

The aim of Mathematica's review was to identify, assess, and rate the efficacy of programs to reduce teen pregnancy, STDs, and associated risky sexual behavior. Studies were selected: (1) from published research syntheses and websites of relevant research and policy organizations, (2) by issuing public calls for studies to solicit new and unpublished research, (3) by conducting keyword searches of numerous electronic databases (e.g., PsycInfo, MedLine, Cochrane, etc.), (4) by searching relevant scientific journals, and (5) by reviewing professional conference proceedings.

Intervention programs were considered for review if they were conducted on US youth with an average age of maximum 19 years or less. There was no lower limit on age for studies to enter the review. Interventions focused on a range of approaches to prevent teen pregnancy, such as encouraging teens to wait to have sex, providing information on contraception, teaching refusal skills, or discussing health consequences of sexual activity. Studies were included if they were conducted or published since 1989. In order to be included, studies had to measure one of the following risky sexual behaviors or outcomes: sexual activity (i.e., initiation, frequency, and number of partners), contraceptive use, incidence of STDs, and pregnancies or births.

Studies that met the criteria were each assessed for quality of research design and implementation. The highest study quality rating was assigned to randomized controlled trials and studies that randomly assigned subjects to groups. Quasi-experimental designs with an external comparison group received a moderate rating. The moderate rating was also applied to random assignment designs that did not meet other criteria for the highest rating (i.e., attrition or reassignment). Finally, studies that did not meet these criteria received low rating. A total of 289 studies met the review screening criteria and were assessed for methodological quality.

Final Study Sample

Out of the 289 studies reviewed by Mathematica we selected the studies that (1) received either high or moderate methodological quality rating and (2) assessed STD incidence either via self-report or a test. A total of 20 interventions met these criteria. One intervention could not be included in the analysis because the information available was insufficient to calculate the effect size.[13] One article reported the effect of two interventions compared to that of a control group. Because of the assumptions of independence of

---

[13] The authors were contacted but could not be reached.

meta-analysis (Viechtbauer, 2010), only one intervention was included. We chose the intervention with richer content. This resulted in a final sample of 18 studies (Figure 1, marked with an asterisk in the reference list).

2.2. Study Information

Methodological quality was controlled for by considering ratings assigned by Mathematica. Most interventions had multiple follow up measurements. To measure effects at similar time intervals, incidence of STDs in the first (last) follow up was considered for long (short) studies.[14] Properties of studies identified as potential moderators of effects (e.g., Johnson et al., 2011; Kirby et al., 2007) were independently coded by the two authors of this paper. These properties included (1) *characteristics of the studies* (outlet, year of publication, dependent measure(s), number of follow-up months, and setting), (2) *characteristics of the sample of participants* (age at the onset of the intervention, gender composition, and if specific population was targeted), (3) *characteristics of the intervention* (type: abstinence vs. comprehensive[15], length, parent involvement, and control group content), (4) *content of the intervention* (knowledge about STDs[16], communication and negotiation skills, condom use skills, sex refusal skills, and gifts/vouchers offered), and (5) *intervention effects on sexual behaviors* (reduction in sexual activity, increase in condom use, and reduction of number of partners in intervention relative to control).

Due to the limited information available for some interventions or the variety of measures used, quantification of some characteristics was either

---

[14] Studies were considered to be short (long) when they lasted 36 months or less (more than 36 months).

[15] Abstinence interventions promoted abstinence from sexual activity until marriage and excluded discussion of the use of contraceptives. Comprehensive interventions described at least one contraceptive method and/or teaching skills related to safer sexual behavior.

[16] Interventions were classified as including STD knowledge if at least one of the following topics was discussed: nature of STDs, how to recognize or prevent them, and how to reduce the risk of contracting an STD.

impossible or not meaningful (e.g., length of intervention as it would range from a few minutes to years). We resolved this by assigning categorical ratings when they were meaningful. For variables where disagreement between coders was possible, inter-rated reliability was calculated using the Krippendorff's alpha, which is appropriate for both continuous and categorical outcomes (Hayes & Krippendorf, 2007). Reliability ranged from 0.66 to 1.00 with an average of 0.90 (*SD*=0.13). Disagreements were resolved through discussion.

2.3. Measures and Analyses

From the selected articles we pooled the number of detected STDs in the intervention and control groups and the respective sample sizes. These numbers were then transformed into a log relative risk (logRR; i.e., the log of the risk in the treated group divided by the risk in the control group (Aldreson & Green, 2002)). A negative logRR value reflects risk reduction after the intervention. The logRR was used as measure of effect size for the meta-analysis because it makes effect sizes symmetric around 0 and their distribution closer to normal (Viechtbauer, 2010). For the sake of clarity, we also computed the risk difference (RD; i.e., the risk in the intervention group minus the risk in the control group (Aldreson & Green, 2002)). Here a negative value implies a reduction of risk in the intervention group (e.g., RD=–.03 means absolute risk reduction of 3% in the intervention relative to control group). We also computed relative risk reduction (RRR; i.e., the STD rate in the control group minus the STD rate in intervention group, divided by the STD rate in the control group). We fitted fixed-, random-, and mixed-effects models using Restricted Maximum Likelihood Estimation in the *R* package *metafor* following Viechtbauer (2010). Studies were weighed using the standard "inverse-variance" method. In particular, in the fixed effects model, studies were weighed by the inverse of their sampling variances. In the random effects model, studies were weighed by the inverse of the sum of the sampling variances and the residual heterogeneity [19]. Each potential moderator of intervention efficacy was examined separately because the low number of studies did not allow testing multiple moderators simultaneously. In order to reduce the probability of Type I error in the large number of moderator tests we set the alpha level to 0.025.

3. Results

In the following, we first briefly describe the main characteristics of the 18 interventions. First, we examined the correlations between the potential moderators of the intervention effect. We computed simple correlations between moderators as the low number of studies did not permit multiple moderator tests. Simple correlations, however, can be informative and help detect characteristics that overlap between certain types of interventions (e.g., abstinence interventions not teaching condom skills). Second, we assessed the average effectiveness of interventions in a fixed and a random effects model. Finally, we examined the effect of moderators on intervention effectiveness (mixed model).

3.1. Description of Interventions

The studies included in the meta-analysis sampled a total of 15,579 adolescents with an average age of 14.40 years (*SD*=2.71) at the start of the intervention. Interventions either targeted a mixed population of boys and girls (56%), or were specifically designed for girls (44%). Some interventions (39%) targeted specific populations like African-Americans, Latinos, or both. The number of months at which follow up took place ranged from 2 to 108 (*M*=31.11, *SD*=27.98). Sixty-one percent of the studies were published in scientific journals, whereas 49% were unpublished reports. The majority of the studies had a high quality rating (78%), assessed STDs via self-report measures (78%), and offered multiple intervention sessions (83%). Length of intervention sessions ranged from a one-time 20 min. session to daily 2.5 hour after school sessions, which students could attend for up to four years.

The programs were conducted at school (56%) or at a clinic (39%), and one intervention (5%) took place in a marine training camp. No interventions included institutionalized adolescents or vulnerable groups (e.g., drug users).

Thirty-nine percent of the interventions were based on the premises of abstinence education. The rest (61%) were classified as comprehensive (i.e., they aimed at improving skills and promoting safe sexual practices). Interventions were mostly theory-based, informed by social cognitive theory (Bandura, 1996), theory of reason action (Marcoux & Shope, 1997), the health-belief model (Janz & Becker, 1984), and others. Fifty percent of the studies included a control group which received some sexual education content (e.g., sexual education at the school). Besides STD incidence, all studies measured condom use. A majority also assessed effects on frequency of sexual activity (61%) and on the number of sexual partners (72%). All interventions either had no effect or improved the respective behavior. No intervention showed an undesirable effect on risky sexual behaviors.

Overlapping moderators. All unpublished studies were abstinence programs, while all published studies were comprehensive education programs. All abstinence programs took place in a school setting and assessed STDs with self-report measures. No abstinence program taught condom use skills or had an effect on any of the preceding behavioral outcomes. However, these variables still varied between the comprehensive programs. Hence, we decided to conduct a moderator analysis on all programs, and verify if the effect still persists when only comprehensive programs are included.

Studies targeting only females tended to take place in a clinic setting, $\varphi_c = .78$, $p = .004$, and use laboratory tests as opposed to self-reported STD incidence, $\varphi_c = .60$, $p = .011$. Studies conducted in a clinic tended to use laboratory tests, $\varphi_c = .67$, $p = .018$. Most of the interventions providing STD knowledge also taught condom use and communication and negotiation skills but did not teach sex refusal skills, all φ_c between .50 and .59, $p<.05$. Most interventions that did not teach refusal skills were comprehensive programs, φ_c =.65, p=.006. The three sexual behavior outcomes we examined (i.e., condom use, frequency of sexual intercourse, and number of partners) were also interrelated, φ_c between .62 and .71, $p<.03$. Interventions which had an effect on one outcome tended to have an effect on the others as well.

3.2. Average effectiveness of interventions

Effect sizes in risk difference (RD) and their variabilities are plotted in Figure 1. First, we created a funnel plot of the study effect sizes to investigate bias in publication (Figure 2). The test for funnel plot asymmetry was not significant, t= .94, df=16, p=.36, and the funnel plot looked fairly symmetrical. The only exception was one outlying study with a small sample size, which found a large intervention effect (Prado et al., 2007). These results indicated that the logRR was an appropriate measure of effect size for this sample and that there was no sign of publication bias. The test for heterogeneity did not indicate significant variability among the true effect sizes, $Q(17) = 23.78$, p=.125, suggesting that the fixed-effects model is appropriate.

To assess the average effectiveness of the conducted interventions, we first fitted a fixed-effects model. The fixed-effects model indicated that interventions reliably reduced STD incidence, $logRR$=–.16, 95% CIs excluding 0 (–.26 to –.05), p=.004. On average interventions reduced the risk of contracting an STD by 1percentage point (mean RD, Figure 1; i.e., from 7 to 6 of 100 people). In other words, interventions reduced STD incidence by 17% (RRR).

To test if intervention success can be inferred beyond this sample of studies, we ran the more conservative random-effects model. The effect in this model was not reliable with $logRR$=–.12, 95% CIs including 0 (–.27 to .03), p=.13, I^2=27.97%, suggesting that we cannot infer an effect for a random set of studies not included in the sample considered in this analysis (Hedges & Vevea, 1998).

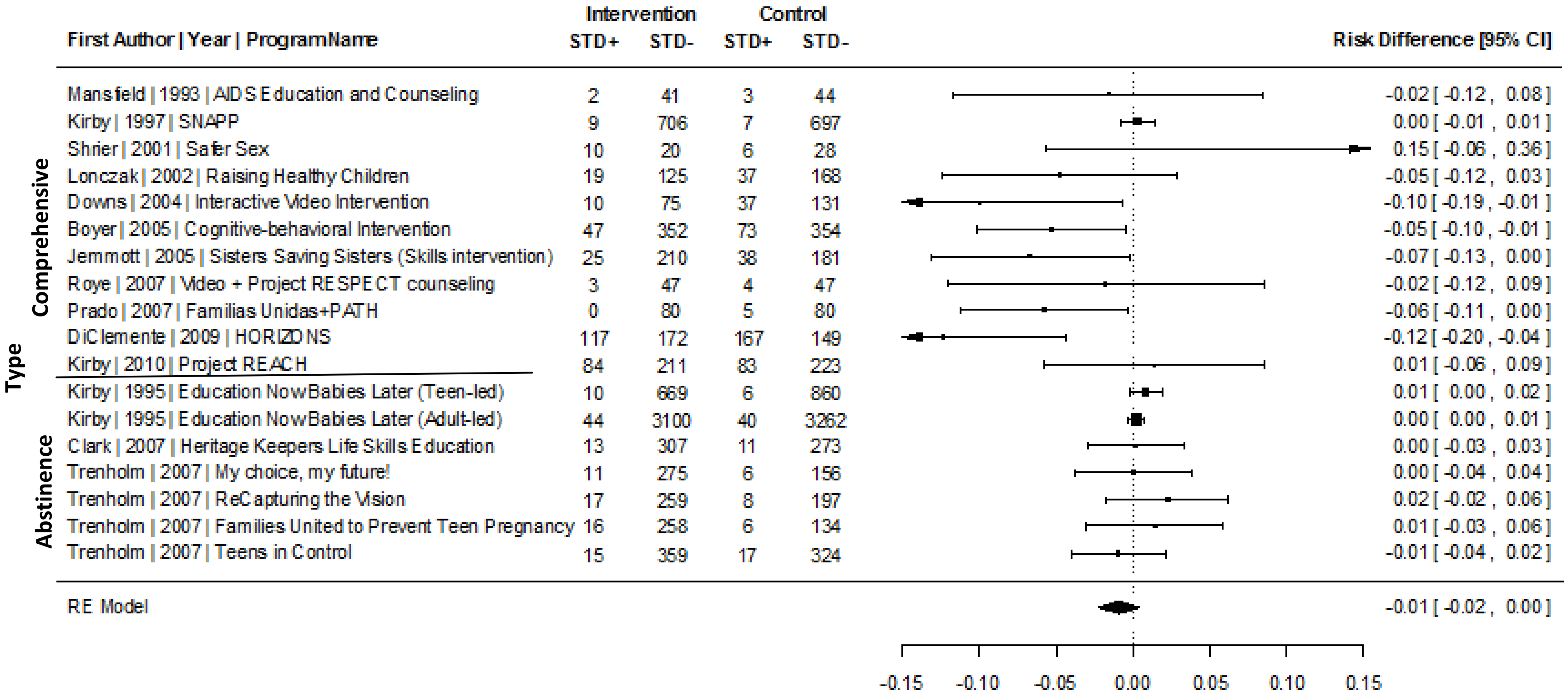
On average interventions were only modestly successful. However, Figure 1 shows that some interventions achieved a more substantial reduction of STD rates than others. In the following section, we examined what factors accounted for the success of these interventions.

3.3. Which interventions were successful?

Detailed statistical results of each moderator test can be found in Table 1.

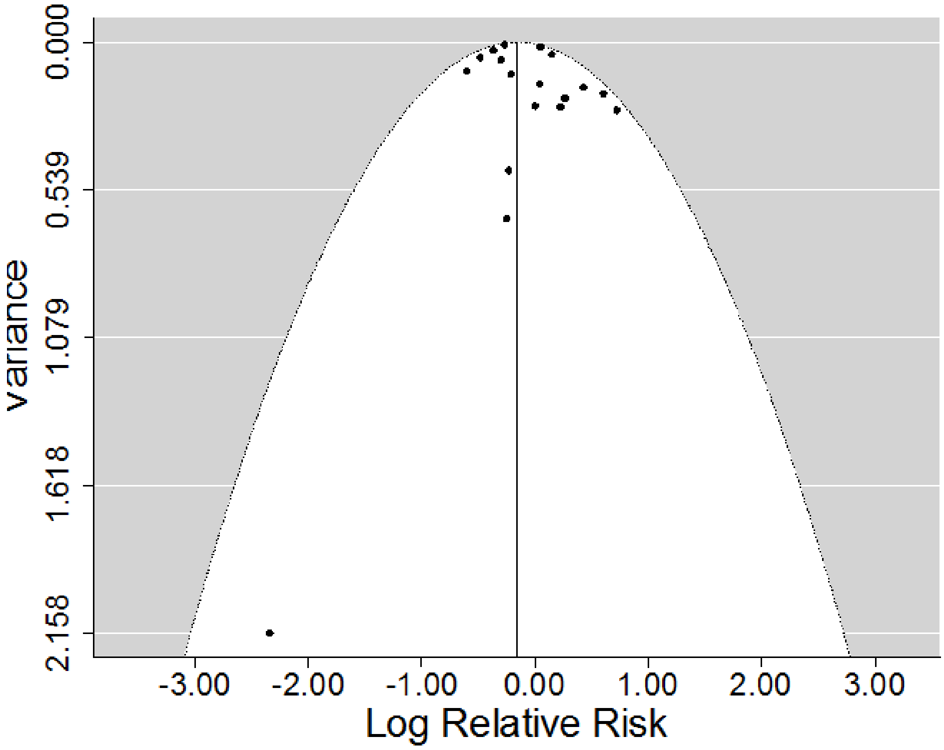
Study, sample, and intervention characteristics. Abstinence interventions did not reduce STD incidence. In contrast, interventions offering comprehensive education did by an average of 4 percentage points (23% RRR). The type of dependent measure had an effect: Interventions that measured STDs with a laboratory test showed a reduction of STD incidence. However, studies that used self-reported measures of STD incidence did not show an effect. Due to the overlap between dependent measures and type of intervention (i.e., all abstinence programs used self-report measures), we ran additional analysis and confirmed that the effect persisted also among comprehensive programs only (p<.01).

Figure 1. Raw data and summary of study effect sizes by type (abstinence vs. comprehensive). The plotted measure is risk difference (intervention-control). A negative value implies risk reduction in the intervention relative to the control group. Error bars around the effect size reflect within study variability. CI=confidence intervals. STD+ =number of people with a STD; STD– = number of people without a STD.



Type	First Author | Year | Program Name	Intervention STD+	Intervention STD-	Control STD+	Control STD-	Risk Difference [95% CI]
Comprehensive	Mansfield | 1993 | AIDS Education and Counseling	2	41	3	44	-0.02 [-0.12 , 0.08]
	Kirby | 1997 | SNAPP	9	706	7	697	0.00 [-0.01 , 0.01]
	Shrier | 2001 | Safer Sex	10	20	6	28	0.15 [-0.06 , 0.36]
	Lonczak | 2002 | Raising Healthy Children	19	125	37	168	-0.05 [-0.12 , 0.03]
	Downs | 2004 | Interactive Video Intervention	10	75	37	131	-0.10 [-0.19 , -0.01]
	Boyer | 2005 | Cognitive-behavioral Intervention	47	352	73	354	-0.05 [-0.10 , -0.01]
	Jemmott | 2005 | Sisters Saving Sisters (Skills intervention)	25	210	38	181	-0.07 [-0.13 , 0.00]
	Roye | 2007 | Video + Project RESPECT counseling	3	47	4	47	-0.02 [-0.12 , 0.09]
	Prado | 2007 | Familias Unidas+PATH	0	80	5	80	-0.06 [-0.11 , 0.00]
	DiClemente | 2009 | HORIZONS	117	172	167	149	-0.12 [-0.20 , -0.04]
	Kirby | 2010 | Project REACH	84	211	83	223	0.01 [-0.06 , 0.09]
Abstinence	Kirby | 1995 | Education Now Babies Later (Teen-led)	10	669	6	860	0.01 [0.00 , 0.02]
	Kirby | 1995 | Education Now Babies Later (Adult-led)	44	3100	40	3262	0.00 [0.00 , 0.01]
	Clark | 2007 | Heritage Keepers Life Skills Education	13	307	11	273	0.00 [-0.03 , 0.03]
	Trenholm | 2007 | My choice, my future!	11	275	6	156	0.00 [-0.04 , 0.04]
	Trenholm | 2007 | ReCapturing the Vision	17	259	8	197	0.02 [-0.02 , 0.06]
	Trenholm | 2007 | Families United to Prevent Teen Pregnancy	16	258	6	134	0.01 [-0.03 , 0.06]
	Trenholm | 2007 | Teens in Control	15	359	17	324	-0.01 [-0.04 , 0.02]
	RE Model					-0.01 [-0.02 , 0.00]

Figure 2. Funnel plot of effect sizes for the fixed effects model, *K*=18. The logRR measure of effect size is plotted against a measure of study precision (variance). One would expect that less precise studies show more variable effect sizes.



Additionally, the methodological quality rating was marginally significant, QM test= .952, *p*= 0.084. Studies with high quality rating reported reduced rates of STDs, *logRR*=–.19, 95% CIs (–.36 to –.01), *p*=.034, while studies with moderate rating reported no difference, *logRR*=.12, 95% CIs (–.22 to .46), *p*=.483. The year of publication, average age of participants at baseline, setting (school or clinic), gender composition of the sample, number of months at which follow up was recorded, the content received by the control group (unrelated or at least some degree of sexual education), and whether a special population was targeted (e.g., Latinas or African Americans) had no influence on the effect size.

Intervention components. Interventions that taught communication and negotiation skills or condom use skills decreased STD incidence by 30% (RRR). Interventions that did *not* teach sex refusal skills decreased STD incidence by 28% (RRR). Provision of STD knowledge had no significant impact on STD incidence, QM test=4.58, *p*=.101. However, dividing the sample according to provision of STD knowledge showed that studies teaching STD knowledge reduced incidence of STDs, *logRR*=–.19, 95% CIs (–.36 to –.01), *p*=.036, while those not teaching STD knowledge had no effect, *logRR*=.07, 95% CIs (–.24 to .38), *p*=.432. These effects remained for comprehensive programs only, *p*s<.04. The participation of parents and the provision of condoms or vouchers after the intervention had no effect.

Intervention effects on sexual behaviors. Several reports did not assess if the interventions had an effect on frequency of sexual activity or the number of sexual partners. Overall, these interventions were successful at reducing STD incidence. Among studies which did measure these outcomes, the interventions which (1) reduced sexual activity (28% RRR), (2) reduced the number of sexual partners (36% RRR), or (3) increased condom use significantly reduced STD incidence (28% RRR). These effects remained significant in analyses on comprehensive programs only (*p*≤.05).

4. Discussion

This meta-analysis summarized evidence accumulated in the past twenty years from interventions aimed to reduce risky sexual behavior in US adolescents. We analyzed interventions with a broad spectrum of aims like social development, delay of sexual initiation, and prevention of HIV, STDs, and unwanted pregnancy among others. Most previous reviews focused on examining the effects of interventions on adolescents' attitudes, behavioral intentions, and sexual behavior (Garcia-Retamero & Cokely, 2014; Kirby & Laris, 2009; Kirby et al., 2007; Silva, 2002). In contrast, the current analysis focused on health outcomes, namely incidence of STDs.

Table 1. Effect sizes at each level of the value of moderator variables. K=number of studies. RD=risk difference. logRR = log value of the relative risk. SE=standard error of the mean. CI-L/H = confidence interval lower/higher bound for logRR. *Note*: *One study was omitted because it was unclear if this education was provided.

		Moderator test						Effects at each level of the moderator variables			
	Moderator	QM test	P-value	K	Value	RD	logRR	SE	P-value	CI-L	CI-H
Study Characteristics	Type	8.85	.012	11	Comprehensive	–.04	–.22	.08	.006	–.38	–.06
				7	Abstinence	.00	.16	.14	.263	–.12	.45
	Dependent Measure	20.09	<.0001	14	Self-report	.00	.06	.08	.453	–.10	.23
				4	Test	–.07	–.32	.07	<.0001	–.47	–.18
Intervention Components	Communication & negotiation skills	18.04	.0001	7*	No	.00	.11	.10	.303	–.09	.31
				10	Yes	–.03	–.28	.07	<.0001	–.41	–.14
	Condom use skills	16.55	.0003	11	No	.00	.04	.08	.647	.13	.21
				7	Yes	–.04	–.29	.07	<.0001	–.43	–.15
	Sex refusal skills	8.21	.0165	12	No	–.03	–.21	.08	.008	–.37	–.06
				6	Yes	.00	.16	.15	.279	–.13	.46
Intervention Effects on Sexual Behaviors	Reduced sexual activity	10.22	.0167	7	Not assessed	–.05	–.20	.09	.025	–.37	–.03
				9	No	.00	.15	.14	.258	–.11	.43
				2	Yes	–.07	–.43	.22	.047	–.87	–.004
	Increased condom use	7.86	.0196	13	No	.00	–.01	.09	.936	–.18	.17
				5	Yes	–.06	–.33	.11	.005	–.56	–.09
	Reduced the number of partners	9.76	.0208	5	Not assessed	–.05	–.19	.10	.062	–.40	.01
				9	No	.00	.15	.14	.257	–.11	.42
				4	Yes	–.05	–.31	.14	.025	–.56	–.04
Overall: Fixed effects model				18		–.01	–.16	.06	.004	–.26	–.05

Additionally, to the best of our knowledge, this is the first meta-analysis examining potential moderators of programs' effect on rates of STDs. We included 18 interventions which assessed STD rates at follow-up, had sound methodology, and provided enough data to compute effect sizes.

Overall, interventions reduced rates of STD by 17%. In absolute terms they achieved a risk reduction of 1 percentage point: without intervention 7 out of 100 adolescents contracted an STD; participating in an intervention reduced this number to 6 out of 100. This effect was significant in the fixed effects model. However, allowing for variability between effects in a more conservative random effects model rendered this effect unreliable. This result suggests that our conclusions cannot be generalized to studies not considered in the current meta-analysis (Hedges & Vevea, 1998). The average intervention effect was smaller than that reported in a recent meta-analysis on the efficacy of interventions to reduce risk of HIV (Johnson et al., 2011). Importantly, Johnson et al.'s (2011) analysis focused on interventions which had at least some HIV-related content and therefore excluded some abstinence and pregnancy programs. In contrast, we included these programs in our analysis, which contributed by decreasing the overall effect. Indeed, interventions based on abstinence education had no effect on STD rates while interventions which were classified as comprehensive reduced STD rates by 23%. This finding is consistent with previous reviews reporting that abstinence education has little or no effect on sexual behaviors like abstinence from sexual activity (Silva, 2002)), delay of initiation of sex (Kirby, 2008), or frequency of sexual intercourse (Johnson et al., 2011). Abstinence education does not provide knowledge or skills related to STD prevention (e.g., condom use skills), which adolescents need once they are sexually active. Thus, abstinence programs can even have a negative effect on rates of STD. However, this is not what evidence suggests.

There were further important moderators of intervention effectiveness. To illustrate, the four interventions using a STD laboratory test as dependent measure detected a 7 percentage point risk reduction in the intervention group. In contrast, studies using self-reported STD incidence did not detect any difference. This result suggests that measuring STD incidence with self-report might be problematic. Reporting a STD diagnosis might be associated with social desirability biases or lack of knowledge among adolescents. As other authors did (DiClemente et al., 2008), we recommend assessing STD incidence among adolescents with laboratory tests when possible. Similarly, only studies of high as opposed to moderate methodological quality showed intervention effects. This again emphasizes the role of successful execution (e.g., low attrition) and good methodology in properly evaluating intervention effects, especially when the outcome is rare.

Interventions which taught communication and negotiation skills or condom use skills reduced STD incidence by 30% (see Johnson et al., 2003, 2011, for similar results). Additionally, successful interventions provided STD related knowledge. However, this effect was not as reliable as the effect of the skill trainings. Importantly, intervention reducing STD rates did not teach sex refusal skills. We think that sex refusal skills can reduce STDs by delaying sexual initiation or reducing frequency of intercourse, an effect similar to that of general abstinence education. This might not be optimal as it might fail to equip adolescents with the necessary skills to protect themselves from diseases and unwanted pregnancy once they do become sexually active. Under such conditions sex refusal skills might be futile in STD prevention. Indeed, interventions which did not provide sex refusal skills tended to be comprehensive rather than abstinence-based and provided extensive STD knowledge. This can explain the presence of an effect when there was no training in sex refusal skills.

We showed that interventions with an effect on relevant behavioral outcomes reduced incidence of STDs by around 30%. Namely, interventions which succeeded in reducing the frequency of sexual intercourse and the number of partners, as well as increasing condom use, resulted in lower STD incidence. This finding shows that all these behaviors can be realistic pathways to STD prevention. Comparing the effects of the three behaviors on STD rates and testing which one has stronger influence can be informative. For example, frequency of sexual intercourse might not have a strong effect on STDs if condom use is consistent. However, due to the currently low number of studies this comparison was not feasible. Future research can investigate this issue once more evidence accumulates.

To conclude, we examined a variety of interventions administered in the past twenty years among US adolescents with the general purpose of reducing risky sexual behaviors. We estimated that these interventions reduced STD rates by only 1 percentage point (17% RRR). However, there were significant differences in intervention success based on program components. Interventions offering comprehensive STD education in the form of prevention-related skills had more success and reduced STD incidence by around 3 to 4 percentage points (23% RRR). These results suggest that future interventions aiming to reduce incidence of STDs should offer STD knowledge and solid communication and condom use skills to adolescents. Based on our analysis we estimate that properly designed interventions can decrease STD incidence by around 30%.

5. References

Alderson, P., & Green, S. (2002). *Open learning material for reviewers*. Cochrane Collaboration. Accessed on 6 August 2013 from http://www.cochrane-net.org/openlearning/HTML/mod0.htm

Bandura A. (1996). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall 1996.

*Boyer, C. B., Shafer, M. A., Shaffer, R. A., Brodine, S. K., Pollack, L. M., Betsinger, K., ... & Schachter, J. (2005). Evaluation of a cognitive–behavioral, group, randomized controlled intervention trial to prevent sexually transmitted infections and unintended pregnancies in young women. *Preventive Medicine*, *40*(4), 420-431.

Centers for Disease Control and Prevention. (2013). *Incidence, prevalence, and cost of sexually transmitted infections in the United States.* Factsheet. Accessed on August 6 2013 from http://www.cdc.gov/std/stats/STI-Estimates-Fact-Sheet-Feb-2013.pdf

*Clark, M. A., Trenholm, C., Devaney, B., Wheeler, J., & Quay, L. (2007). *Impacts of the Heritage Keepers life skills education component.* Princeton, NJ: Mathematica Policy Research Inc.

DiCenso, A., Guyatt, G., Willan, A., & Griffith, L. (2002). Interventions to reduce unintended pregnancies among adolescents: systematic review of randomised controlled trials. *BMJ*, *324*(7351), 1426.

DiClemente, R. J., Crittenden, C. P., Rose, E., Sales, J. M., Wingood, G. M., Crosby, R. A., & Salazar, L. F. (2008). Psychosocial predictors of HIV-associated sexual behaviors and the efficacy of prevention interventions in adolescents at-risk for HIV infection: what works and what doesn't work? *Psychosomatic Medicine*, *70*(5), 598-605.

*DiClemente, R. J., Wingood, G. M., Harrington, K. F., Lang, D. L., Davies, S. L., Hook III, E. W., ... & Hardin, J. W. (2004). Efficacy of an HIV prevention intervention for African American adolescent girls: a randomized controlled trial. *JAMA*, *292*(2), 171-179.

*DiClemente, R. J., Wingood, G. M., Rose, E. S., Sales, J. M., Lang, D. L., Caliendo, A. M., ... & Crosby, R. A. (2009). Efficacy of sexually transmitted disease/human immunodeficiency virus sexual risk–reduction intervention for African American adolescent females seeking sexual health services: A randomized controlled trial. *Archives of Pediatrics & Adolescent Medicine*,*163*(12), 1112-1121.

*Downs, J. S., Murray, P. J., de Bruin, W. B., Penrose, J., Palmgren, C., & Fischhoff, B. (2004). Interactive video behavioral intervention to reduce adolescent females' STD risk: A randomized controlled trial. *Social Science & Medicine*, *59*(8), 1561-1572.

Forhan, S. E., Gottlieb, S. L., Sternberg, M. R., Xu, F., Datta, S. D., McQuillan, G. M., ... & Markowitz, L. E. (2009). Prevalence of sexually transmitted infections among female adolescents aged 14 to 19 in the United States. *Pediatrics*, *124*(6), 1505-1512.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied*, *17*(3), 270-287.

Garcia-Retamero, R., & Cokely, E. T. (2014). The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making*, *27*(2), 179-189.

Garcia-Retamero, R., & Cokely, E. T. (2015). Simple but powerful health messages for increasing condom use in young adults. *The Journal of Sex Research*, *52*(1), 30-42.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*,*1*(1), 77-89.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486-504.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education & Behavior*, *11*(1), 1-47.

*Jemmott, J. B., Jemmott, L. S., Braverman, P. K., & Fong, G. T. (2005). HIV/STD risk reduction interventions for African American and Latino adolescent girls at an adolescent medicine clinic: a randomized controlled trial. *Archives of Pediatrics & Adolescent Medicine*, *159*(5), 440-449.

Johnson, B. T., Carey, M. P., Marsh, K. L., Levin, K. D., & Scott-Sheldon, L. A. (2003). Interventions to reduce sexual risk for the human immunodeficiency virus in adolescents, 1985-2000: a research synthesis. *Archives of Pediatrics & Adolescent Medicine*, *157*(4), 381-388.

Johnson, B. T., Scott-Sheldon, L. A., Huedo-Medina, T. B., & Carey, M. P. (2011). Interventions to reduce sexual risk for human immunodeficiency virus in adolescents: a meta-analysis of trials, 1985-2008. *Archives of Pediatrics & Adolescent Medicine*, *165*(1), 77-84.

Kirby, D. B. (2008). The impact of abstinence and comprehensive sex and STD/HIV education programs on adolescent sexual behavior. *Sexuality Research & Social Policy*, *5*(3), 18-27.

*Kirby, D., Korpi, M., Adivi, C., & Weissman, J. (1997). An impact evaluation of project SNAPP: an AIDS and pregnancy prevention middle school program. *AIDS Education and Prevention*, *9*, 44–61.

*Kirby, D., Korpi, M., Barth, R.P, *et al*. (1995). *Evaluation of Education Now and Babies Later (ENABL): final report*. Berkeley, CA: Family Welfare Research Group.

Kirby, D., & Laris, B. A. (2009). Effective curriculum-based sex and STD/HIV education programs for adolescents. *Child Development Perspectives*, *3*(1), 21-29.

Kirby, D. B., Laris, B. A., & Rolleri, L. A. (2007). Sex and HIV education programs: their impact on sexual behaviors of young people throughout the world. *Journal of Adolescent Health*, *40*(3), 206-217.

*Kirby, D., Raine, T., Thrush, G., Yuen, C., Sokoloff, A., & Potter, S. C. (2010). Impact of an intervention to improve contraceptive use through follow-up phone calls to female adolescent clinic patients. *Perspectives on Sexual and Reproductive Health*, *42*(4), 251-257.

*Lonczak, H. S., Abbott, R. D., Hawkins, J. D., Kosterman, R., & Catalano, R. F. (2002). Effects of the Seattle Social Development Project on sexual behavior, pregnancy, birth, and sexually transmitted disease outcomes by age 21 years. *Archives of Pediatrics & Adolescent Medicine*, *156*(5), 438-447.

*Mansfield, C. J., Conroy, M. E., Emans, S. J., & Woods, E. R. (1993). A pilot study of AIDS education and counseling of high-risk adolescents in an office setting. *Journal of Adolescent Health*, *14*(2), 115-119.

Marcoux, B. C., & Shope, J. T. (1997). Application of the theory of planned behavior to adolescent use and misuse of alcohol. *Health Education Research*, *12*(3), 323-331.

Mathematica Policy Research & Child Trends. (2012). *Identifying programs that impact teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors: a review protocol.* US Department of Health and Human Services: Office of Adolescent Health. Accesed on August 6 2012 from http://www.hhs.gov/ash/oah/oah-initiatives/teen_pregnancy/db/eb-programs-review-v2.pdf

Picot, J., Shepherd, J., Kavanagh, J., Cooper, K., Harden, A., Barnett-Page, E., ... & Frampton, G. K. (2012). Behavioural interventions for the prevention of sexually transmitted infections in young people aged 13–19 years: a systematic review. *Health Education Research*, *27*(3), 495-512.

*Prado, G., Pantin, H., Briones, E., Schwartz, S. J., Feaster, D., Huang, S., ... & Szapocznik, J. (2007). A randomized controlled trial of a parent-centered intervention in preventing substance use and HIV risk behaviors in Hispanic adolescents. *Journal of Consulting and Clinical Psychology*, *75*(6), 914-926.

*Roye, C., Silverman, P. P., & Krauss, B. (2007). A brief, low-cost, theory-based intervention to promote dual method use by black and Latina female adolescents: a randomized clinical trial. *Health Education & Behavior*, *34*(4), 608-621.

Sales, J. M., Milhausen, R. R., & DiClemente, R. J. (2006). A decade in review: building on the experiences of past adolescent STI/HIV interventions to optimise future prevention efforts. *Sexually Transmitted Infections*, *82*(6), 431-436.

*Shrier, L. A., Ancheta, R., Goodman, E., Chiou, V. M., Lyden, M. R., & Emans, S. J. (2001). Randomized controlled trial of a safer sex intervention for high-risk adolescent girls. *Archives of Pediatrics & Adolescent Medicine*,*155*(1), 73-79.

Silva, M. (2002). The effectiveness of school-based sex education programs in the promotion of abstinent behavior: a meta-analysis. *Health Education Research*, *17*(4), 471-481.

*Trenholm, C., Devaney, B., Fortson, K., Quay, L., Wheeler, J., & Clark, M. (2007). *Impacts of four Title V, Section 510 abstinence education programs: final report.* Princeton, NJ: Mathematica Policy Research Inc.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package*. Journal of Statistical Software*, *36*, 1–48.

Wildsmith, E., Schelar, E., Peterson, K., & Manlove, J. (2010), *Sexually transmitted diseases among young adults: prevalence, perceived risk, and risk-taking behaviors*. Washington, D.C.: Child Trends 2010.

CHAPTER 10:
Experiences with HPV vaccine communication

The content of this chapter has been published as Petrova, D., Gray Brunton, C., Jaeger, M., Lenneis, A., Muñoz, R., Garcia-Retamero, R., Todorova, I. (2015). The views of young women on HPV vaccine communication in four European countries. *Current HIV Research*, *13*(5), 347-358.

Chapter 10

The views of young women on HPV vaccine communication in four European countries

The Human Papilloma Virus (HPV) is the most common sexually transmitted infection (STI) and can cause cervical cancer. Two vaccines are available to protect against the most common strands of the virus. Vaccination programs differ across Europe but most neglect young adults, who are the group with the highest risk of contracting STIs. Our aim was to explore the views of young women about the HPV vaccine communication strategy from four European countries: Scotland, Spain, Serbia, and Bulgaria. These countries are characterized by different cervical cancer prevalence and vaccine implementation policies. We conducted focus group discussions with young women (aged 18-26) with various vaccination histories in a purposive sample. We subjected the data to thematic analysis with the purpose of identifying themes related to communication about the HPV vaccine. We recorded the information sources mentioned by participants. Participants discussed numerous sources of vaccine-related information. They approached information critically rather than naively and questioned the sources' trustworthiness and motives. Participants desired transparent information about the risks of the virus and the risks and benefits of the vaccine. These risks and benefits were individualized in view of personal and external factors. Particular aspects of the vaccine and the way information was communicated resulted in feelings of uncertainty. There were notable cross-cultural differences in experiences with HPV vaccine communication. Our results suggest that transparent risk communication about the HPV vaccine is valued by young women. In addition, both individual and culturally-dependent factors influenced experiences with, and preference for information.

1. Introduction

Effective risk communication is fundamental to improve health behavior in general (Lipkus, 2007; Weinstein, 1990) and to prevent sexually transmitted infections (STIs) in particular. The Human Papillomavirus (HPV) is the most common STI and may cause a number of serious health conditions, including genital warts and cervical cancer (Centers for Disease Control, 2013). Cervical cancer is among the most prevalent cancer types in women worldwide. In 2005, almost 260.000 women were estimated to have died of cervical cancer with 80% of the cases occurring in the developing world (World Health Organization, 2007). Since 2006, two brands of the HPV vaccine have been introduced to guard against those HPV strains responsible for 70% of cervical cancer cases (European Cervical Cancer Association (ECCA), 2009). The HPV vaccine is effective in preventing HPV infection and HPV related diseases in 90%–100% of cases (La Torre, de Waure, Chiaradia, Mannocci, & Ricciardi, 2007). It is well tolerated with mild side effects, which makes it one of the most efficacious vaccines available (Bonanni et al., 2011).

The HPV vaccine is increasingly available in a number of countries across the world (Hopkins & Wood, 2013). In Europe it has been introduced with different implementation policies ranging from school-based mass immunization programs to on-demand delivery or private sector provision only. This has led to different uptake rates. For example, a school-based program in Scotland resulted in 92% coverage, while free on-demand provision in Greece in only 9% coverage (ECCA, 2009). Overall most countries in North-Western and Central Europe have implemented some sort of organized vaccination program or campaign (Bonanni et al., 2011). In contrast, vaccination efforts in some South-Eastern European countries are still developing despite the region's higher burden of cervical cancer (Arbyn et al., 2007; Seme et al., 2012). Generally, programs have focused on vaccinating preadolescent girls before the onset of sexual activity. Although a few countries offer catch-up programs for women in their twenties, this age group has mostly been neglected by vaccine policies (Bonanni et al., 2011), possibly due to cost-effectiveness (Canfell et al., 2012). However, vaccination of young women is generally recommended as young women can still benefit from the vaccine, even if they have been infected with HPV. Although HPV vaccination has no therapeutic efficacy in women who have already been infected, infection with one virus type does not impede vaccine-induced protection from another type (Schiller, Castellsagué, & Garland, 2012). Few countries offer catch-up programs for women in their twenties, and even fewer provide full insurance coverage (Bonanni et al., 2011). Most significantly, women in their twenties show the highest HPV infection prevalence, which makes vaccination particularly beneficial (Dunne et al., 2007).

Besides organized vaccination efforts, public communication is essential for informed decision making. The introduction of a new vaccine is accompanied by the challenge of educating the target audience with clear information (Sherris et al., 2006). Next to implementation policies, communication source, content, and form of the information shapes beliefs and behavior related to the HPV vaccine (see also Bigman, Cappella, & Hornik, 2010; Gerend & Shepherd, 2012; Lipkus, 2007; Nan & Madden, 2012). Concerns about risks and side effects can decrease intentions to get vaccinated (Kester, Zimet, Fortenberry, Kahn, & Shew, 2013; Ritov & Baron, 1990). Even though the vaccine is not reported to have severe side effects, concerns about its risks may potentially be fuelled by negative reports in the media (Gainforth & Latimer, 2012). Anti-vaccine activism on the Internet and social media that advances the view that pharmaceutical companies exert undue influence for their own benefit can further increase skepticism (Franco et al., 2012).

Uncertainties are also increased by what is perceived by users as a lack of transparency in communication. Transparent communication may require complete, non-misleading, and factual information about the HPV vaccine and its associated risks in public media (Bodemer, Müller, Okan, Garcia-Retamero, & Neumeyer-Gromen, 2012; Garcia-Retamero & Galesic, 2013). A recent analysis shows that health provider's materials may require more accurate, complete, and consistent information (Steenbeck, MacDonald, Downie, Appleton, & Baylis, 2012). Some information might intentionally be left out, as it is potentially controversial to talk about young girls' sexuality (Mishra & Graham, 2012). A Canadian study, for example, found that around 20% of mothers assumed that HPV vaccination promoted promiscuity and were therefore less likely to get their daughters vaccinated (Ogilvie et al., 2007). This concern can be further fuelled by media coverage on the topic (Forster, Wardle, Stephenson, & Waller, 2010). Even though this issue is expected to be less severe in more liberal societies (e.g., Denmark, see Mortensen, 2010), health providers might still neglect the sexual nature of the transmission in the promotion of the vaccine (Mishra & Grahan, 2012). Generally, there appears to be a need for improved information which includes not only transparent risk and benefit statistics (Garcia-Retamero & Galesic, 2013; Ritov & Baron) but which also addresses the socio-cultural aspects that might influence vaccination decision making.

Scientific reasoning alone might not be sufficient to deal with anti-vaccination attitudes (Franco et al., 2012). The success of health efforts like the HPV vaccine introduction may be influenced by a country's unique history and culture and their implications for people's trust in health care systems, physicians and pharmaceutical companies (e.g., Craciun & Baban, 2012; Todorova, 2011). For example, physicians' recommendations increased HPV vaccination rates among 19–26 years old women in Australia (Rosenthal, Dyseon, Pitts, & Garland, 2007). However, physicians' recommendations might not be taken into account if there is mistrust towards physicians and the domestic health care system such as in Hungary (Marek et al., 2011).

More understanding is needed of contextual aspects including historical experiences, religious or political affiliation, and socioeconomic status that play a central role in shaping public trust in vaccines (Larson, Cooper, Eskola, Katz, & Ratzan, 2011). It has become clear that a "one-size-fits-all-approach" to introduce a HPV program is problematic (Franco et al., 2012). For example, the most prominent constructs used in health behavior models might not be universal and social context and cultural factors can influence health-protective behavior directly (Pasick et al., 2009). Hence, an approach which focuses exclusively on individual level factors like those of most health-behavior models (e.g., Health-Belief Model (Rosenstock, 1974), Theory of Planned Behavior (Aizen, 1991)), can be insufficient to understand vaccination behavior in general and experience with information in particular. Qualitative research approaches can help in addressing this issue. They allow a thorough analysis and exploration of health issues and are valuable because they can take into account individual views embedded in social and cultural surroundings (Lyons, 2011). The differences in the implementation strategies of the vaccine across Europe offers a rich environment to investigate the reception of risk communication and the influence of policy and culture on individual decision making.

The aim of this paper is to explore young women's experiences with risk information about HPV and the HPV vaccine. We explored how women evaluated the quantity and quality of HPV-related information and information sources they consulted, and how this information affected their decisions about the HPV vaccine. We focused on young women between the ages of 18 and 26 in four European countries: Scotland, Spain, Serbia, and Bulgaria. We chose this age group because women aged 18-26 are at high risk of contracting STIs and can benefit from vaccination. However, most HPV vaccination policies and campaigns target younger girls, thereby excluding this vulnerable age group. We chose these countries because they have different implementation policies, thereby offering a variety of information exposure and experiences with HPV information.

2. Methods

We chose qualitative methodology in order to provide a rich description of women’s experiences as embedded in their social, cultural, and personal context. We conducted focus group discussions with university educated women in Scotland, Spain, Serbia, and Bulgaria. These countries are characterized by different HPV vaccination strategies targeting adolescent girls. In particular, Scotland employs a national mass school-based program, Spain offers free on-demand provision, while Serbia and Bulgaria offered only private provision at the time of data collection (ECCA, 2009; Seme et al., 2012). None of the countries had implemented a catch-up vaccination program or any opportunity for vaccine cost reimbursement for women aged 18-26 (Bonanni et al., 2011; Seme et al., 2012). As a reimbursement for participation, participants entered a raffle with a maximum prize of 30 GBP in Scotland, 30 EUR in Spain, 3000 RSD in Serbia, and 50 BGN in Bulgaria, respectively. Ethics approval was obtained from the Research Ethics and Governance Committee at the leading institution Edinburgh Napier University, as well as the respective institutions at the University of Granada, the Serbian Psychological Society, and Sofia University.

2.1. Participants

Participants were women between 18 and 26 years old. They were recruited through purposive and snowballing sampling via posters, e-mailing lists, and forums in the local universities. We attempted to include women with various experiences and vaccination histories, ranging from being vaccinated in school in the past to never having heard of the HPV vaccine before. We conducted eleven focus group discussions with 54 young women in total. Table 1 describes the demographic characteristics of the participants and their allocation among focus groups. All participants were university students, representing a mix of majors.

2.2. Data collection

The discussions were conducted between December 2011 and July 2012. The same procedure was followed in all four countries. All materials were translated into the local language. Participants were informed about the nature of the focus group discussion, signed an informed consent, and filled in an anonymous questionnaire with demographic information, sexual and vaccination history. All discussions lasted around one hour and were facilitated by two native speakers according to the same topic guide. The topics covered included participants' opinions and experiences with vaccination in general and the HPV virus and vaccine in particular, including sources of information about the HPV vaccine. Towards the end of the discussions participants were presented with a sample of publicly available brochures about the HPV vaccine. These always included a brochure from the respective country in the local language. Additionally, brochures from the participating countries as well as other European countries were available. As the latter were in a foreign language, they were provided simply for visual inspection. The purpose of presenting the brochures was to provide a further basis for discussion. Participants were given time to get acquainted with the materials and had the opportunity to comment.

2.3. Analysis

The discussions were digitally recorded and transcribed by native speakers of the local language. Personally identifying information was removed. The Spanish, Serbian and Bulgarian transcripts were translated into English by native speakers who were fluent in English. To ensure quality of the translation random subset back-translations were performed [39].

Table 1. Participant demographic characteristics

	Scotland	Spain	Bulgaria	Serbia	Total
Number of focus groups	3	3	3	2	11
Number of participants	10	25	10	9	54
Mean age	22.2	19.8	22.3	22.8	21.8
Sexually active	9	19	10	8	46
Vaccinated	4	3	0	0	7
Prior knowledge about HPV vaccine	10	21	9	3	43
Attitude towards the HPV vaccine					
Interested	1	8	1	2	12
Unsure	6	8	7	7	28
Not interested	3	9	2	0	14

Thematic analysis. The transcripts were subjected to thematic analysis following Braun and Clarke (Brislin, 1990). Thematic analysis is a method for identifying, analysing, and reporting recurring patterns (i.e., themes) within data. We adopted a mixture of inductive and theoretical approaches to our thematic analysis. In particular, we coded the data specifically for issues related to risk communication (as opposed to all interesting issues that might be present). However, we chose no particular theory or model against which to evaluate the data but preferred an exploratory approach. The topics we coded for included but were not limited to the following: (1) perceived risk of HPV and cervical cancer, (2) experience with, and perceived quality, of risk information, (3) preferred information sources, (4) experience with, and perceived quality, of information sources, and (5) preferred/recommended manner of risk communication. We followed the steps outlined by Braun and Clarke (2006) while taking into account the cross-cultural nature of our data. In particular, each coder (authors DP, AL, and RM) first familiarized herself with the data after receiving training in qualitative research. Then each transcript was coded for significant features by two coders. These initial codes were then organized into initial themes by each individual coder. Next, each coder was assigned a country for which she compared, summarized, and organized the initial themes into a thematic map first for each focus group separately and then for the respective country as a whole. The similarities and differences between the country's thematic maps were organized into an overall thematic map for the whole data set following extensive discussion between all coders. The final thematic map was refined by the first author after "going back to the data" to check if the final themes fit in relation to the coded abstracts (e.g., identifying telling quotes to illustrate the themes).

Information sources analysis. Participants discussed numerous sources of HPV-related information. Hence, we decided to subject our qualitative data to quantitative analysis (e.g., seeCokely & Kelley, 2009) in order to give a more thorough overview of women's experiences. We recorded the frequency with which participants mentioned 1) a unique information source from which they *obtained information* about the HPV vaccine[17], and 2) a unique information source they *would consult* if they required such information. We grouped the sources into meaningful categories and computed the number of times a source from each category was mentioned in each discussion.

3. Results

3.1. Thematic analysis

We identified two main themes: **Critical Appraisal of Information** and **Risk Adjustment** (Table2). **Critical Appraisal of Information** encompassed two stages of obtaining HPV-related knowledge. The first stage of information reception was more passive: it reflected participants' appraisals of the available information. Common subthemes were *perceived lack of*

---

[17] No differentiation was made between instances in which the participant herself searched for information from a source or was approached by the information source as this was not specified on many occasions.

information and *role of the media*. The subtheme *transparency* marked the transition to a more active stage of knowledge acquisition: Participants evaluated the transparency of the communicated evidence and purposefully searched for what, in their perspective, was transparent and trustworthy information. Other common subthemes were *active information search* and *information sources and reliability*. The theme **Risk Adjustment** reflected the stage of decision making about the HPV vaccine. The subtheme *my risk, my benefits* captured how participants perceived and adjusted their risk of HPV and the benefit of the vaccine in view of a number of personal factors. Other subthemes were *alternatives to vaccination* and *uncertainty,* which reflected how risks and benefits regarding HPV and the vaccine were interpreted in view of contextual factors. These themes were common across the four countries. However, some of the themes were more prominent in certain countries than others. Below we clarify the themes in more detail and provide excerpts from the focus group discussions to illustrate the participants' perspectives.

Table 2. Thematic framework developed from the analysis.

Critical Appraisal	Lack of information Role of the media Transparency Active search for information Information sources and reliability
Risk Adjustment	My risk, my benefit Alternatives to vaccination Uncertainty

Critical Appraisal of Information

This theme illustrates participants' overall critical approach towards the information they encountered. Most young women in this study felt insufficiently informed and desired transparent information about the risks and benefits of the virus and the vaccine. When assessing the credibility of information, they questioned the trustworthiness and motives of information providers. Sometimes this led them to feel supported, and sometimes misled or manipulated.

***We don't know*: Lack of information.** Participants' knowledge about HPV ranged from little to more advanced knowledge. However, a common perception was that the obtained information was insufficient to make an informed decision about vaccination. This idea was shared both by participants who took part in a school vaccination campaign and participants who had barely heard of the HPV vaccine. Further, this perceived lack of information often translated into an inference that the vaccine was not very important. Some participants thought that if the vaccine was indeed beneficial, then more information would have been distributed, reflecting the support of authorities for the vaccine.

"And if the vaccine is still so important to cervical cancer, why aren't they still giving information about it, so that more girls can get the jab? It was a craze that has stalled now." (P5, FG2, Spain)[18]

"Well, for me the primary reason is that, in reality, if the government wanted to support this thing, it would have provided more information, as a first step towards action." (P1, FG1, Bulgaria)

Young women presented themselves as part of an information generation with access to knowledge sources. Yet, this privilege stood in contrast to their lack of knowledge about HPV. This is illustrated by a

---

[18] Quotes identifiers are (Participant number, Focus Group Number, Country).

participant from Serbia who reflected on the far greater information disadvantages experienced by people with fewer educational and socio-economic opportunities:

"Well you [have us here] who are relatively educated and informed, who are following all sorts of things and then if we do not know ... Then I cannot imagine someone who does not have access to the internet and no access to television, how can someone know? But that's what we have here..." (P4, FG2, Serbia)

***You shouldn't be advertising*: The role of the media.** Women often talked about the information they had obtained from the media such as television and newspapers. In particular, television advertisements were frequently discussed. Their strength was seen in sparking interest and raising awareness. In addition, women recognized that the media played a major role in how information was dissipated, and "that the image of vaccines depends mostly on how the media plays it" (P5, FG1, Spain). As a consequence one needed to rely on one's own active critical skills rather than be a passive recipient of advertising or news coverage. For example, the quote below illustrates how the information provided through advertising was seen as shallow and insufficient, which implied that one had to go beyond such superficial information to be truly informed.

"There was a lot of advertisements on TV, there was a craze, a lot of advertising on TV selling it but actually about information, they didn't inform at all, [they told you] it just prevented the possibility of having cervical cancer but they didn't tell you the cons or pros that the decision of getting the vaccine or not could have for you." (P5, FG2, Spain)

Often the media were seen as "selling" the vaccine rather than providing information about it and this had a negative connotation. This idea was reinforced by presenting the vaccine in absolutely positive terms and in a commercial context. A common perception was that important medicines like the HPV vaccine were potentially life-saving and should not be sold akin to a commercial product by using emotions or an attractive appearance.

"I don't think that the brochure must have a very attractive appearance, because its aim is not to sell a product, but mainly to inform." (P1, FG1, Bulgaria)

Participants often "saw through" the advertiser's intentions or recognized their influential techniques, labelling them as "manipulative" or "directed". The overall negative perception of the commercialized vaccine is illustrated by the following discussion of an advertisement among Scottish participants:

"P4: *I think it was like "arm against cervical cancer" or something (voices overlapping: yeah)...I think that's what it was called, there was a chain of girls holding arms and it was all sorts of girls you know... just all sorts of people you know (laughing)...and I think that was quite appealing to the age group.* ***P3:*** *I think it was quite directed...which then to me I think that's wrong, you shouldn't be advertising, you know thinking about the audience like that because it's not a product, it's er.. medical, you're not trying to sell this!" (FG2, Scotland)*

It's here to convince**: Transparency.** Participants often commented that the information they encountered was unbalanced. What they meant was that it was presented purposefully in a biased fashion, or certain aspects were omitted, in order for the information to be more persuasive. Such information was evaluated unfavourably, while information which consisted of presenting both pros and cons was evaluated more favourably and as more trustworthy. Among the omissions mentioned were mainly the possible side effects of vaccination, or the fact that there are more strains of the virus than the vaccine protects you from. In that sense, presenting the vaccine as "100% protection" was seen as misleading.

[Commenting on a brochure] "Also this one, it seems that it's here to convince that it's necessary to get the vaccine, let's say. It should also say

different points of view, the side effects and all that. I see that it mostly supports a position." (P2, FG1, Spain)

"I liked very much in that brochure that it said everything very realistically, and not to delude you, for example, that if you get vaccinated you will 100% be... It is directly described that they don't guarantee." (P3, FG2, Bulgaria)

When discussing what sort of information women wanted to receive about the virus and the vaccine again great emphasis was placed on transparency. This meant presenting facts in an open, unbiased fashion, as well as tailored to the needs of the recipients, such that it was understandable and adequate. For example, one woman expressed her desire to see more statistical information as opposed to persuasive appeals, or "fairy tales" as she referred to them with irony.

"Scientific, statistics, something very exact which can answer your questions in a more detailed way, not like a fairytale. Because brochures are like fairytales. This is great, absolutely, you will reinvent the wheel if you get the vaccine." (P3, FG3, Bulgaria)

***Find out for yourself*: Active information search**. Women's perception was that they would not receive the necessary information about the vaccine unless they actively searched for it. Rather, they had to be pro-active and inform themselves.

"I think that there is information if you look for it, you know, if you have gone to get some information. If you haven't gone to get information, no..." (P2, FG1, Spain)

Some of them reported "doing their own research" about the vaccine as the information they initially received was insufficient. On the one hand, it was perceived as routine to look for additional information, as not all the necessary knowledge could be conveyed in one source or be targeted to all sorts of recipients. However, some young women perceived that information was withheld, so that they would have to make an effort to get it. For example, one Bulgarian participant commented on her impression that Bulgarian doctors were reluctant to provide information unless you specifically asked about it:

"Now, to be honest, doctors here are like that: unless you ask about every single detail that you are concerned about, they will not start explaining and give you the information you need. You have to ask and find out for yourself. So it is questionable how much doctors are informing us." (P2, FG1, Bulgaria)

***We first investigate the doctor*: Information sources and their reliability.** Women discussed numerous sources of information, ranging from health professionals to people with experience with vaccination (see information sources analysis for further details). Among these, family and friends played a supportive rather than an informative role, except for when they were health professionals. Rumours and personal stories of people who got vaccinated were also pervasive, especially in the Spanish focus groups. Most significantly, gaining a second opinion was emphasised in the discussions. . In that sense, sources needed to be "triangulated" in order for the information to be verified. One example suggested by participants was searching for information on the Internet and then talking to one's doctor.

Most importantly, information was not assimilated blindly. Rather, participants would evaluate the source's trustworthiness and motives before making a decision based on the provided information. This included seeking "expert" opinion from someone who was seen as competent rather than trusting random posts on the Internet. It was recognized that even expert sources could be biased for various reasons (e.g., doctors having a contract with pharmaceutical companies). This was particularly the case in Serbian and Bulgarian discussions where health professionals were often mistrusted. One solution to this problem was to consult a doctor who had a personal connection to the family.

"I rely mainly on the opinions of relatives and family friends who are doctors or of similar profession, what should and what should not be necessary, which are bad and good sides of the vaccine." (P2, FG1, Serbia)

"Well, we first investigate that doctor and see if he's good. If he has a lot of patients...I, by the way, have investigated my personal gynaecologist through the Internet and I know that he is very good, so maybe I would trust him." (P2, FG2, Bulgaria)

Risk Adjustment

This theme captures how participants acted on the information they had obtained and approached the decision about HPV vaccination. Very prominent in this theme was the tendency for risks and benefits to be strongly individualized in view of personal, generational, and societal factors. Certain aspects of the vaccine and the available information resulted in feelings of uncertainty. The core issues in this theme were mostly similar across countries. However, some of them were manifested in different ways in the Western and Eastern discussions. We elaborate on these cross-cultural differences below.

***It depends*: My risk, my benefits**. When considering their decision regarding the HPV vaccine, women tended to talk about the risks of contracting HPV and the associated benefit of the vaccine in personal rather than in absolute terms. Risk perceptions were adjusted according to factors like age, sexual history and current relationship status. As women were mainly in their twenties, they considered the potential number of past sexual partners as indicative of the likelihood that they had already encountered the virus. This led to further discussions around doubts of the vaccine's effectiveness for their age group.

"Now I'm thinking well it's quite useless cos I've already had some sexual partners and I could have it already and I think smear tests are useful for that reason." (P1, FG1, Scotland)

"Well, we are already sexually active, and you are supposed to get the vaccine before you become active, because then it is most....I forgot the word... most effective. Actually, at this point, you can't be sure about the effectiveness of the vaccine." (P2, FG1, Bulgaria)

Young women also considered the number of potential future sexual partners as indicative of the risk group they belonged to. In that sense, having one stable partner was seen as a means of risk reduction which rendered vaccination at that point less beneficial. This is illustrated by one Serbian participant who was considering what she would have done if she were offered the vaccine. She considered HPV vaccination strictly as "a matter of choice" because not everyone was at the same risk of infection; risk was individualized and dependent on lifestyle:

"First, you need to start with introducing some educational measures and then later to make it simple: a matter of choice. It depends on what kind of person you are. I do not know, maybe I wouldn't have agreed, because I just know who I am, I do not let myself into relationships, how to put it... I do not have many partners." (P3, FG1, Serbia)

***What's the use of it?* Alternatives to vaccination.** Women discussed alternative ways of protection for themselves, for the female population and society in general. In Spain and Scotland these were most often cervical screening with smear-tests and regular visits to the gynaecologist. Questions were often raised about the relationship between screening and vaccination, as well as possible "removal" of the virus if discovered through screening. For example, one Spanish participant thought that the vaccine was useless if one had regular check-ups.

"Screening is something that might be...even more effective, then if you keep track you're going to avoid it and all that. Had I known it earlier, we would have saved a lot of money from vaccines [laughter]. If I'm going to have my check, what's the use of it?" (P3, FG2, Spain)

Sexual education and responsible sexual behaviour were seen as alternative means of protection for the young women themselves and for the population as a whole. This perspective was emphasized in Bulgaria and Serbia. Participants discussed ignorance about HPV and the HPV vaccine in the context of their societies: they identified the lack of sex education from an early age as the reason for the lack of awareness of sexually-transmitted infections. Education was seen as a way to promote responsible sexual behaviour, which in turn would reduce risks. However, for participants mass HPV vaccination was at an advanced stage of protection, which was still far away in view of the two countries' lack of policies and lack of public awareness regarding STIs.

The joint influence of personal and societal factors in considering the benefits of vaccination against other alternatives is illustrated by two Bulgarian women:

*"**P1:** [Reading off a brochure, turns to P3] Even if you get vaccinated, it is no guarantee and you have to continue with the screening. **P3:** Then what is the point of this vaccine? **P1:** So, 450 + 75 leva*[19] *every month... **P3:** This is meant for a country whose health system is waaay ahead of ours." (FG3, Bulgaria)*

***We can't know*: Uncertainty**. Women's decisions regarding the vaccine were characterized by three dimensions of uncertainty. One concerned their personal judgment of the available information such as knowing whom to trust or what information is most important. This was very much related to the subtheme about information sources and their reliability, and its impact on decisions. The uncertainty relating to finding a competent and trustworthy information source is illustrated by a heated discussion among Spanish participants about the role of health professionals in providing information:

---

[19] ~EUR265

*"**P5:** Nowadays you go to the doctor, or to the pharmacist, asking them about the human papillomavirus or about the vaccine and for sure they still don't know what to tell you. **P3:** And it's not only that, because if you go to the doctor you go to a place where you feel safe. **P4:** That's why I'm saying that I, if the doctor tells me about getting it and I get it because I trust him and that's it, but if I, for example, am not really sure because I might think, why is it that they give it to a type of person and the other…? You know? I can ask myself a thousand questions, then if I'm interested, because it's my health, I can also go to other, to other kind…**P3:** But it's supposed that these questions should be answered by the doctor? **P4:** It's supposed, you've said it yourself, it's supposed! **P5:** Come on, if the doctor doesn't answer them, where do you think you'd go?" (FG2, Spain)*

Another type of uncertainty referred to the vaccine's side effects. In the Spanish and Scottish focus groups discussions of media reports about deaths or personal stories of people experiencing violent side effects were a frequent topic. While some participants disregarded these reports as "hyped up" or eventually discredited, for others they undermined the vaccine's credibility or made them reconsider the vaccine's benefits.

"I think that there are many side effects, which aren't known. Only fever, dizziness, fainting are known. That there are a lot of others, that there are people who get paralysed or can even die and all that, and not a lot is known." (P9, FG3, Spain)

"I go for it, because…I didn't know the risks with a vaccine. I wasn't really aware of them, but I figured side effects I can put up with it if the rest is beneficial…"(P2, FG2, Scotland)

Finally the vaccine's novelty brought about feelings of uncertainty related to future consequences, especially among Western-European participants. Young women wanted vaccines that were 100% safe, well-tested, and approved. At the same time they recognized that the HPV vaccine had only been in use for a few years and that its long-term effects had not been established. On the one hand, this made some young women feel akin to “guinea pigs” and reject the vaccine. On the other hand, it was recognized that one generation had to be the first to take the risk for the greater final good. Some participants were under the impression that even experts had doubts, and this contributed to their uncertainty:

“But yeah it’s too new to say whether it is good or not. There should be further studies. The results of what is happening now should be released.”(P1, FG1, Scotland)

“Yes, yes, they [HPV vaccines] are a part of the immunization schedule already, and I think it’s rushed, that they should have researched more. Knowing that it isn’t, that the vaccine isn’t developed, knowing that you don’t know what can happen when you get it, putting it up on the schedule... And there are many doctors, the medical opinion has been divided because of that, because it’s been a big debate, and I think it’s very important that the medical opinion has been divided.” (P5, FG1, Spain)

3.2. Information source analysis

Participants brought up information sources on 134 occasions, with some participants discussing multiple sources and some discussing none. The information sources were grouped into 8 categories (see Table 3). Table 4 shows the number of times women discussed having obtained information from each information source. Table 5 shows the number of times women discussed that they would prefer to consult each information source. On 66 (49%) instances young women discussed sources from which they obtained HPV-related information and on 68 (51%) instances sources from which they would prefer to obtain such information. Participants most often reported having received information from the media, family/friends, and health professionals. Interestingly, in the Spanish focus groups the most frequent source was family/friends. In the Serbian discussions women mentioned obtaining information from few sources overall. The most preferred source across all countries was health professionals, followed by searching on the Internet. The third category, surprisingly, was family or friends who are also health professionals, prominent in the Serbian focus groups. It is noteworthy that scientific sources emerged as a preferred category that participants would consult in three out of four countries.

Table 3. Information sources, descriptions, and frequency with which they were discussed.

Source	Description	Frequency	Percent
Health professionals	Doctor, general practitioner, gynecologist, nurse, etc.	40	29.9
Internet	"The internet", Google, forums, general search without specification	25	18.7
Media	TV, news, newspapers, entertainment culture , TV commercials, advertisements	21	15.6
Family/friends	Mothers, friends, acquaintances, etc.	18	13.4
Family/friends who are health professionals	Mother who is a doctor, cousin who is a nurse, etc.	12	9.0
School	"In school" in general, listened to a talk at school	7	5.2
Scientific sources	Medical publications, scientists	7	5.2
People with experience	Vaccinated individuals, people who have decided about vaccination	4	3.0
Total		134	100.0

Table 4. The number of times women in each country discussed having obtained information from each information source.

	Bulgaria		Scotland		Serbia		Spain		All countries	
Source	Count	%	Count	%	Count	%	Count	%	Count	%
Health professionals	2	15%	2	13%	3	38%	6	21%	13	20%
Internet	2	15%	2	13%	-	-	1	3%	5	8%
Media	5	38%	6	38%	2	25%	7	24%	20	30%
Family/friends	4	31%	1	6%	1	13%	10	34%	16	24%
Family/friends health professionals	-	-	2	13%	1	13%	2	7%	5	8%
School	-	-	2	13%	1	13%	3	10%	6	9%
Scientific sources	-	-	1	6%	-	-	-	-	1	2%
People with experience	-	-	-	-	-	-	-	-	-	-
Total	13	100%	16	100%	8	100%	29	100%	66	100%

Table 5. The number of times women in each country discussed that they would prefer to consult each information source.

	Bulgaria		Scotland		Serbia		Spain		All countries	
Source	Count	%	Count	%	Count	%	Count	%	Count	%
Health professionals	6	38%	4	29%	9	38%	8	57%	27	40%
Internet	7	44%	5	36%	5	21%	3	21%	20	29%
Media	-	-	-	-	1	4%	-	-	1	1%
Family/friends	1	6%	-	-	1	4%	-	-	2	3%
Family/friends health professionals	-	-	2	14%	5	21%	-	-	7	10%
School	-	-	-	-	-	-	1	7%	1	1%
Scientific sources	2	13%	3	21%	1	4%	-	-	6	9%
People with experience	-	-	-	-	2	8%	2	14%	4	6%
Total	16	100%	14	100%	24	100%	14	100%	68	100%

4. Discussion

The goal of this paper was to explore the experiences of young women aged 18-26 with communication about the HPV vaccine. This included information and information sources about HPV and the HPV vaccine. Being well-informed about the risks and benefits associated with HPV and the vaccine is not only ethically desirable but also the first step towards vaccine uptake, as several health-behavior models suggest (Fisher, 2012). Qualitative, cross-cultural methodology allowed us to examine in depth the interplay between micro (individual) and macro level (societal) factors related to obtaining information about HPV (Fisher, 2012). Our participants were university educated women from four European countries with different vaccine implementation policies. This age group is especially interesting for two reasons. First, most vaccination policies focus on younger girls (11 to 14 years) and populations beyond this age are not eligible for mass immunization. However, older girls and young women could still benefit from vaccination (Schiller et al., 2012). Second, this specific age group (18 to 26 years) offered insight into a variety of experiences. These included women who were vaccinated at an earlier age in a school vaccination program (Scotland) or via on-demand provision (Spain). Alongside this were women who had relatively less exposure to HPV campaigns (Bulgaria) or almost none (Serbia) (Seme et al., 2012). Hence, a variety of social contexts (e.g., presence of a vaccination program), personal experiences (with information and vaccination), and opinions (pro, against or undecided) were represented in the discussions.

We identified several similar issues in women’s discussions, which we grouped into two common themes. The first theme concerned **Critical Appraisal** of the available information. This referred to both the perceived quantity (sufficient or not) and quality evaluations (transparent or biased) of information. Overall, women perceived that they did not have sufficient knowledge to make an informed decision. Significantly, even women from countries where the HPV vaccine had received more state and media attention felt uninformed regarding essential aspects of the HPV vaccine. Problems with accessing information were particularly striking in Serbia, where the majority of women had no previous information about the vaccine (Table 1). Although we cannot conclude that our sample is representative of the general population, this finding reflects the state of HPV vaccine implementation in Serbia, where vaccination has never been officially recommended and initial campaigns have ceased [9]. In fact, this is the first study to our knowledge to address psychological issues related to HPV in any population in Serbia.

Although we did not assess women's knowledge per se but rather focused on their perception of the degree to which they were informed, we should note that there were common misconceptions, doubts, and questions about the vaccine. These often related to the relationship between sexual activity and the vaccine's efficacy, the degree of risk reduction provided by the vaccine, and the necessity and benefits of screening in combination with vaccination. It might be useful in subsequent research to focus on identifying such questions, so that they can be properly addressed in information materials.

Our analysis of information sources revealed that information about HPV reached young women mainly through the common channels of social transmission (e.g., family and friends), advertisement, and health professionals. However, both the frequency and thematic analyses highlight differences between actual experience and preference for information sources. The most preferred information source was a health professional. This was, however, followed by an Internet search, which often included skimming through "everything that is out there" and selecting reliable information. This again highlights women's preference for self-reliance and the ability to critically examine information. However, it is not clear how well young women can actually identify high quality information on the Internet, or to what degree low quality information influences their decisions. Both the thematic and frequency analysis suggest the potential benefit of directing young women to neutral and credible sources where they can inform themselves in more detail. Future research should investigate what these sources could be depending on the context and how such a strategy influences decision making about the vaccine. For example, an international source provided by authorities perceived as trustworthy (e.g., on a European rather than national level) may be more effective in increasing trust in information. This may be particularly relevant in Eastern Europe where information provided on a local level might be met with scepticism. One potentially useful source in our opinion is the webpage of the European Cervical Cancer Association (http://www.ecca.info/) which offers the same essential information in different languages. Future research can also investigate whether presenting the information about the vaccine using different information formats might improve risk comprehension and risk communication (Garcia-Retamero & Cokely, 2013; Garcia-Retamero ^ Hoffrage, 2013).

Overall women in our sample were critical rather than naïve towards the information they encountered. Participants actively searched for information and preferred transparency: to them this meant presenting all facts (both pros and cons) in an unbiased fashion so that one could make an informed decision. In order to assess the transparency of information, young women considered two types of cues: the information provider's motives (are they genuinely interested in my well-being or are they trying to make money?) and the presence of both positive and negative information (both benefits and side-effects). Previous qualitative studies with young women focused mainly on identifying barriers and facilitators of vaccine acceptance (e.g., Hopfer & Clippard, 2011; Mortensen, 2010). Our thematic analysis with a focus on the HPV vaccine communication strategy revealed that common facilitators of vaccination like health provider recommendation occur within a context and could easily misfire if information is not perceived as transparent. Further, this can vary not only from person to person but also between contexts. Previous vaccination scares in the UK (Wakefield et al., 1998) for example, or the complex socio-cultural implications of a communist legacy in some Eastern European countries (Todorova, 2011) can influence vaccination efforts. Furthermore, people might become mistrustful of health-care in general or be especially sensitive to trustworthiness cues in health information. A very striking example is the recent HPV campaign failure in post-communist Romania, where mistrust and conspiracy theories appeared to be the main reason that vaccination coverage did not go over 2.5% (Craciun & Baban, 2012). Such occurrences suggest that vaccine communication can be more effective if culturally-sensitive issues are anticipated and addressed. Furthermore, it might be useful to further explore culturally defined notions of information transparency. While our analysis revealed that young women had common notions of transparency, transparency as a concept may be subjective. What is regarded as transparent information might constitute culturally-specific dimensions (for example, including information about particular aspects of a vaccination program might be perceived as more essential in some contexts than in others). How transparency is defined in different contexts requires further qualitative study.

Another noteworthy result was related to the commercial side of the vaccine. The image of the HPV vaccine as a commercial product sold on the market was seen as inappropriate and provoked negativity. The vaccine was regarded as a potentially life-saving product, to which everyone should have access. The high costs of this “product” contrasted with the healthcare systems of the European countries where the discussions took place, where primary care is free or insurance is relatively affordable. Further, this negative view of commercialism might not be restricted only to potential recipients of the vaccine. For instance, Bulgarian medical professionals expressed similar scepticism about the effectiveness of the vaccine and the profit behind it (Dimitrova, Panayotova, Todorova, & Alexandrova-Karamanova, 2012).

The second main theme was **Risk Adjustment**. It captured how the known risks and benefits were adjusted in view of personal (e.g., prospective sexual relations) and societal factors (e.g., availability of screening services). Mainly, the risk of contracting HPV was related to the number of partners. This led some women to conclude that a single or long term partner meant reduced risk. A similar finding with US participants was reported by Hopfer and Clippard (2011): "Responses reflected the false belief, in many cases, that monogamy was protective" (p. 272). Our analysis, however, suggests that young women carefully considered the risks and benefits their own situation entailed. Rather than feeling absolutely protected by monogamy, participants compared their risk of contracting HPV from the present to the risk of someone who would be exposed to more partners. In that sense, the vaccine offered less benefit to them at this age and this benefit had to compete with the numerous vaccination costs. These costs included the monetary cost as well as the psychological costs related to informing oneself properly and running additional risks associated with side effects.

There were some notable differences between countries in discussing alternatives and potential uncertainties related to vaccination. Regular screening as an alternative to vaccination was more widely discussed in Scotland and Spain, while educating the public about health-protective behaviours and STIs in general was more prominent in the Bulgarian and Serbian discussions. Further, issues about vaccine novelty, media stories about dangerous vaccine side effects or rumours about people experiencing them were present in the two Western discussions but not in the discussions from the Eastern European countries. These differences might be rooted in the nature and degree of information exposure in these countries. This was supported by the frequency analysis of information sources, showing that the media and social transmission were more common ways to receive information among Scottish and Spanish participants surpassing health professionals. Further, Scotland and Spain have already witnessed HPV vaccination campaigns. Hence, it is to be expected that young women would be more aware of HPV and screening overall, while the participants from Eastern Europe attributed the absence of information to lack of education and support by the state. Similarly, the media in Scotland and Spain were more likely to reflect on potential HPV vaccination scares, while HPV was less reported in the Bulgarian and Serbian media at the time because of the absence of national policies. In that sense, some issues related to vaccine resistance might not only be specific to cultural context but specific to implementation stage. If this is the case, it might be strategic for authorities to anticipate and counteract exaggerated vaccine resistance with appropriate risk communication. Shortly after this research project was finalized the Bulgarian Ministry of Health adopted a vaccination program targeting teenage girls [9]. It will be interesting to see if as implementation progresses, similar issues to those in Scotland and Spain arise in Bulgarian discussion about HPV.

Overall our results emphasized the multifaceted role of information acquisition and evaluation in vaccination decision making. This is supported by the Information-Motivation-Behavioral Skills Model (IMB) (Fisher, 2012). The IMB model goes beyond traditional motivational models and incorporates the role of vaccine-related information, along motivation and behavioural skills in vaccine uptake. Applied to the context of HPV vaccination, the model states that individuals who are well-informed, motivated to act, and possess the relevant behavioural skills are likely to get vaccinated. The IMB model also incorporates the influence of macro-level factors like ethnicity or vaccine cost. Within the framework of the IMB model, our research focused on the role of HPV vaccine information and explored its interplay with such macro-level factors (e.g., media coverage, socio-cultural context, and trust in authorities). The IMB model can be further adapted to accommodate these and other relevant macro-level factors identified by qualitative inquiry and examine their influence on individual decision making in a quantitative paradigm. The IMB model also emphasizes that skills such as securing funding or negotiating health care provider support are essential for vaccine uptake. Our analysis shows that women utilize similar skills even at the earlier pre-decision stage of information acquisition as captured by the theme **Critical Appraisal**. These might include the ability to search for information and to evaluate the credibility of information sources.

There are a few typical limitations and strengths inherent to this type of methodology that should be noted. First, we relied on volunteers to participate in the discussions, which might have attracted participants with a strong opinion about the HPV vaccine. However, the fact that there were participants who had no prior information about the HPV vaccine in all countries but Scotland speaks against this possibility. Second, some original meanings might have been lost after translating the transcripts into English. In order to minimize this problem, we involved native speakers in the analysis, used multiple coders, and chose an analytical procedure which is not highly sensitive to language use (Braun & Clarke, 2006). On the other hand, one major advantage of qualitative methods is that participants' answers are less constrained by the researcher's perspective. This provides the opportunity to discover in depth issues important to participants that might not have occurred to researchers (e.g., Bensing, Rimondini, & Visser, 2013) and to consider participants' reflections in their broader socio-cultural context.

HPV is the most frequent sexually-transmitted infection worldwide and a cause of cervical cancer (CDC, 2013). The HPV vaccine is potentially one way to reduce cervical cancer incidence but questions and challenges to implementation remain (Franco et al., 2012). Informing the public is an essential and ethically desirable step to prevention. However, recent analyses in Europe (Bodemer et al., 2012), the US (Madden, Nan, Briones, & Waks, 2012), and Canada (Steenbeck et al., 2012) reveal that information about HPV might be suboptimal and inaccurate even when provided by official authorities. It is not yet clear how the public deals with such information. Our qualitative, thematic analysis demonstrated that the young university educated women who participated in our discussions took most information about the HPV vaccine with a grain of salt: They evaluated the quality of information and preferred transparent to selective or biased reporting. This suggests that it would be beneficial to direct young women to credible, high-quality information sources they can consult on their own or encourage consultation with a health professional, who was still regarded as the most preferred information source. However, the question remains as to how less privileged groups deal with information. In addition, young women considered the decision to vaccinate against HPV at their age (18 to 26 years) as highly individualized; they considered their lifestyle and potential alternatives to vaccination; they also recognized the uncertainty related to making a decision in view of (un)regulated communication of evidence and uncertainty inherent with new vaccines. Last, the cultural context influenced women's exposure to information and shaped their perceptions of vaccine relevance and safety.

5. References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211.

Arbyn, M., Primic-Žakelj, M., O Raifu, A., Grce, M., Paraskevaidis, E., Diakomanolis, E., ... & Von Karsa, L. (2007). The burden of cervical cancer in South-East Europe at the beginning of the 21st century. *Collegium Antropologicum*, *31*(2), 7-10.

Bigman, C. A., Cappella, J. N., & Hornik, R. C. (2010). Effective or ineffective: Attribute framing and the human papillomavirus (HPV) vaccine. *Patient Education and Counseling*, *81*, S70-S76.

Bodemer, N., Müller, S. M., Okan, Y., Garcia-Retamero, R., & Neumeyer-Gromen, A. (2012). Do the media provide transparent health information? A cross-cultural comparison of public information about the HPV vaccine. *Vaccine*, *30*(25), 3747-3756.

Bonanni, P., Levi, M., Latham, N. B., Bechini, A., Tiscione, E., Lai, P., ... & Boccalini, S. (2011). An overview on the implementation of HPV vaccination in Europe. *Human Vaccines*, *7*(sup1), 128-135.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185-216.

Canfell, K., Chesson, H., Kulasingam, S. L., Berkhof, J., Diaz, M., & Kim, J. J. (2012). Modeling preventative strategies against human papillomavirus-related disease in developed countries. *Vaccine*, *30*, F157-F167.

Centers for Decease Control. (2013). *Genital HPV Infection – Fact Sheet*. Factsheet available at http://www.cdc.gov/std/hpv/HPV–Factsheet–March–2013.pdf. Accessed 17 August 2013.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*(1), 20-33.

Craciun, C., & Baban, A. (2012). "Who will take the blame?": Understanding the reasons why Romanian mothers decline HPV vaccination for their daughters. *Vaccine*, *30*(48), 6789-6793.

Dimitrova, E., Panayotova, Y., Todorova, I., & Alexandrova-Karamanova, A. (2012). Doctors' and parents' perspectives on communication regarding HPV vaccination in Bulgaria. In: *Health, culture and the human body: epidemiology, ethics and history of medicine, perspectives from Turkey and Central Europe* (pp. 341-364). Betim Center Press.

Dunne, E. F., Unger, E. R., Sternberg, M., McQuillan, G., Swan, D. C., Patel, S. S., & Markowitz, L. E. (2007). Prevalence of HPV infection among females in the United States. *JAMA*, *297*(8), 813-819.

European Cervical Cancer Association. (2009). HPV vaccination across Europe. Factsheet available at http://www.ecca.info/fileadmin/user_upload/HPV_Vaccination/ECCA_HPV_Vaccination_April_2009.pdf. Accessed 5 May 2013.

Fisher, W. A. (2012). Understanding human papillomavirus vaccine uptake. *Vaccine*, *30*, F149-F156.

Forster, A., Wardle, J., Stephenson, J., & Waller, J. (2010). Passport to promiscuity or lifesaver: press coverage of HPV vaccination and risky sexual behavior. *Journal of Health Communication*, *15*(2), 205-217.

Franco, E. L., de Sanjosé, S., Broker, T. R., Stanley, M. A., Chevarie-Davis, M., Isidean, S. D., & Schiffman, M. (2012). Human papillomavirus and cancer prevention: gaps in knowledge and prospects for research, policy, and advocacy. *Vaccine*, *30*, F175-F182.

Gainforth, H. L., & Latimer, A. E. (2012). Risky business: risk information and the moderating effect of message frame and past behaviour on women's perceptions of the human papillomavirus vaccine. *Journal of Health Psychology*, *17*(6), 896-906.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science*, *22*(5), 392-399.

Garcia-Retamero, R., & Galesic, M. (2013). *Transparent communication of health risks: Overcoming cultural differences*. New York, NY: Springer.

Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine*, *83*, 27-33.

Gerend, M. A., & Shepherd, J. E. (2012). Predicting human papillomavirus vaccine uptake in young adult women: Comparing the health belief model and theory of planned behavior. *Annals of Behavioral Medicine*, *44*(2), 171-180.

Gray Brunton, C., Farver, I., Jäger, M., Lenneis, A., Parve, K., Patarcic, D., ... & Todorova, I. (2014). Young women's constructions of the HPV vaccine: a cross-cultural, qualitative study in Scotland, Spain, Serbia and Bulgaria. *International Journal of Behavioral Medicine*, *21*(1), 11-19.

Hopfer, S., & Clippard, J. R. (2011). College women's HPV vaccine decision narratives. *Qualitative Health Research*, *21*, 262–77.

Hopkins, T. G., & Wood, N. (2013). Female human papillomavirus (HPV) vaccination: global uptake and the impact of attitudes. *Vaccine*, *31*, 1673–1679.

Kester, L. M., Zimet, G. D., Fortenberry, J. D., Kahn, J. A., & Shew, M. L. (2013). A national study of HPV vaccination of adolescent girls: rates, predictors, and reasons for non-vaccination. *Maternal and Child Health Journal*, *17*(5), 879-885.

La Torre, G., de Waure, C., Chiaradia, G., Mannocci, A., & Ricciardi, W. (2007). HPV vaccine efficacy in preventing persistent cervical HPV infection: A systematic review and meta–analysis. *Vaccine*, *25*, 8352–8358.

Larson, H. J., Cooper, L. Z., Eskola, J., Katz, S. L., & Ratzan, S. (2011). Addressing the vaccine confidence gap. *The Lancet*, *378*(9790), 526-535.

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Medical Decision Making*, *27*, 696–713.

Lyons, A. C. (2011). Advancing and extending qualitative research in health psychology. *Health Psychology Review*, *5*(1), 1-8.

Madden, K., Nan, X., Briones, R., & Waks, L. (2012). Sorting through search results: a content analysis of HPV vaccine information online. *Vaccine*,*30*(25), 3741-3746.

Marek, E., Dergez, T., Kricskovics, A., Kovacs, K., Rebek-Nagy, G., Gocze, K., ... & Gocze, P. (2011). Difficulties in the prevention of cervical cancer: adults' attitudes towards HPV vaccination 3 years after introducing the vaccine in Hungary. *Vaccine*, *29*(32), 5122-5129.

Mishra, A., & Graham, J. E. (2012). Risk, choice and the 'girl vaccine': Unpacking human papillomavirus (HPV) immunisation. *Health, Risk & Society*, *14*(1), 57-69.

Mortensen, G. L. (2010). Drivers and barriers to acceptance of human-papillomavirus vaccination among young women: a qualitative and quantitative study. *BMC Public Health*, *10*(1), 68-82.

Nan, X., & Madden, K. (2012). HPV vaccine information in the blogosphere: how positive and negative blogs influence vaccine-related risk perceptions, attitudes, and behavioral intentions. *Health Communication*, *27*(8), 829-836.

Ogilvie, G. S., Remple, V. P., Marra, F., McNeil, S. A., Naus, M., Pielak, K. L., ... & Patrick, D. M. (2007). Parental intention to have daughters receive the human papillomavirus vaccine. *Canadian Medical Association Journal*,*177*(12), 1506-1512.

Pasick, R. J., Burke, N. J., Barker, J. C., Joseph, G., Bird, J. A., Otero-Sabogal, R., ... & Washington, P. K. (2009). Behavioral theory in a diverse society: Like a compass on Mars. *Health Education & Behavior*, *36*(5 suppl), 11S-35S.

Ritov, I., & Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioral Decision Making*, *3*(4), 263-277.

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education & Behavior*, *2*(4), 328-335.

Rosenthal, D., Dyson, S., Pitts, M., & Garland, S. (2007). Challenges to accepting a human papilloma virus (HPV) vaccine: a qualitative study of Australian women. *Women & Health*, *45*(2), 59-73.

Schiller, J. T., Castellsagué, X., & Garland, S. M. (2012). A review of clinical trials of human papillomavirus prophylactic vaccines. *Vaccine*, *30*, F123-F138.

Sherris, J., Friedman, A., Wittet, S., Davies, P., Steben, M., & Saraiya, M. (2006). Education, training, and communication for HPV vaccines. *Vaccine*,*24*, S210-S218.

Seme, K., Maver, P. J., Korać, T., Canton, A., Částková, J., Dimitrov, G., ... & Kesić, V. (2012). Current status of human papillomavirus vaccination implementation in central and eastern Europe. *Acta dermatovenerologica Alpina, Pannonica, et Adriatica*, *22*(1), 21-25.

Steenbeek, A., MacDonald, N., Downie, J., Appleton, M., & Baylis, F. (2012). Ill-Informed Consent? A Content Analysis of Physical Risk Disclosure in School-Based HPV Vaccine Programs. *Public Health Nursing*,*29*(1), 71-79.

Todorova, I. (2011). Two decades of social change in Central and Eastern Europe: implications for health. In: Friedman HS, Ed. *The Oxford Handbook of Health Psychology* (pp. 824-845). New York: Oxford University Press.

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., ... & Valentine, A. (1998). RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, *351*(9103), 637-641.

Weinstein, N. D. (1999). What does it mean to understand a risk? Evaluating risk comprehension. *Journal of the National Cancer Institute Monographs*, *25*, 15–20.

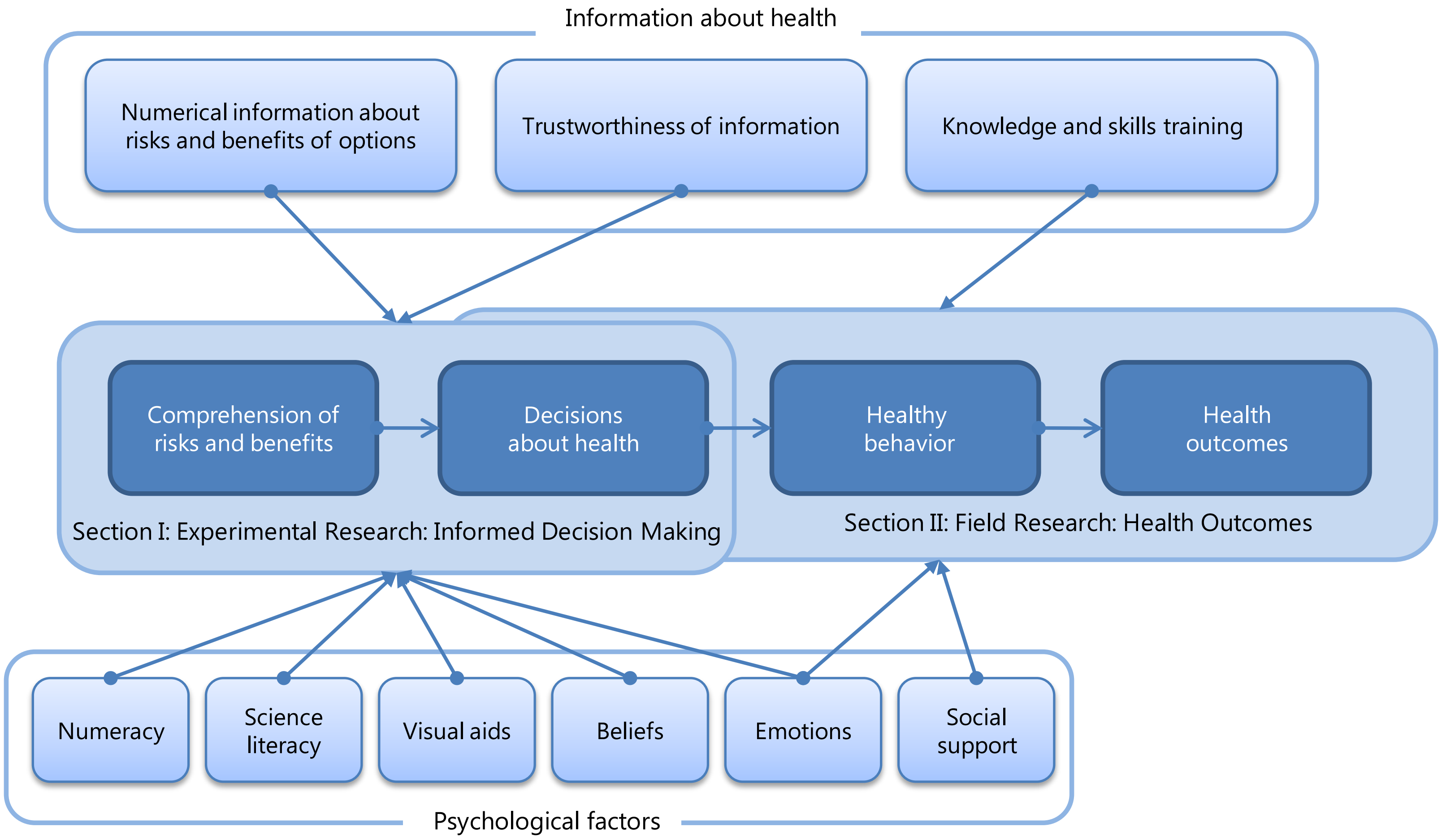
World Health Organization. (2007). *Human papillomavirus and HPV vaccines*. Factsheet available at http://whqlibdoc.who.int/hq/2007/WHO_IVB_07.05_eng.pdf. Accessed 17 August 2013.

DISCUSSION

The aim of this thesis was to contribute to the psychology of informed decision making by investigating the influence of cognitive, social, and environmental factors on decisions about health and health outcomes. We investigated several important psychological factors in two series of studies (see Figure 1). The first series includes experiments testing and estimating the causal influences of health information, cognitive skills, emotions, and beliefs on comprehension of health information and decisions about health. The second series of studies includes field research investigating actual behavior and outcomes in a wide range of patients. This series of studies extends some of the experimental findings beyond the laboratory setting and explores how several additional factors influence health behavior and outcomes directly.

1. Section I: Experimental Research

Informed decision making is recommended for decisions that are preference-sensitive. Patients are encouraged to make decisions considering both the benefits and the risks of different clinical options, as well as their own values and preferences (Betsch et al., 2015; Charles, Gafni, & Whelan, 1999; Fowler, Levin, & Sepucha, 2011; Rimer, Briss, Zeller, Chan, & Woolf, 2004; Salzburg Global Seminar, 2011). Part of this process includes considering evidence-based information often of numeric, probabilistic nature. For this purpose, it is essential to have a suffient level of risk literacy – the ability to understand, evaluate, and make good decisions about risk (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012). This thesis showed that numeracy, science literacy, visual aids, beliefs, and emotions can affect risk literacy and informed decision making directly and/or indirectly via influencing comprehension of numerical health information.


Information about health
Numerical information about risks and benefits of options
Trustworthiness of information
Knowledge and skills training
Comprehension of risks and benefits
Decisions about health
Healthy behavior
Health outcomes
Section I: Experimental Research: Informed Decision Making
Section II: Field Research: Health Outcomes
Numeracy
Science literacy
Visual aids
Beliefs
Emotions
Social support
Psychological factors

1.1. Comprehension of numerical information can promote informed and shared decision making about health.

Understanding statistical evidence is a small yet essential part of the complex decision making process (see Figure 1). Comprehension of numerical risks and benefits associated with different treatment options is both ethically desirable and in certain situations a prerequisite for informed decision making to take place (e.g., Wills & Holmes-Rovner, 2003). Based on a series of experiments in various settings, this thesis shows that comprehension of such evidence has serious implications for the quality of decisions and can thus indirectly affect health. Many patients expect to receive important risk and benefit information and they want this information in a way that is understandable and coming from someone they can trust (Chapter 10). However, information about health is not always transparent, consistent, and reliable (e.g., Bodemer, Müller, Okan, Garcia-Retamero, & Neumeyer-Gromen, 2012; Moynihan et al., 2000). Misunderstanding of risk information can lead to unfounded fear and worry, overprotective behavior, and support for suboptimal public policies that can cost lives (Chapter 4). It can also lead to the adoption of treatments that, although well-intended, are not effective and can cause serious harms (Chapters 1-3). In contrast, better understanding can help fine tune strong emotional responses and correct exaggerated perceptions of risk or benefit (Chapter 1).

Unfortunately, misunderstanding is common among patients and even medical professionals are not immune to the same problems (Chapter 3). Many well-intentioned physicians are prone to misunderstanding risks, resulting in communicating inaccurate risk information to their patients, potentially causing a cascade of serious errors (Chapter 3). In contrast, comprehension tends to improve high-stakes decisions, and often promotes shared decision making (Chapters 1 to 4). Intuitively, physicians who understand the evidence are more likely to provide complete and transparent risk communication to their patients (Chapter 3). In turn, patients who understand the evidence are more likely to want to be involved in the decision making process and share decision making with their physician (Chapter 1). Previous research has linked stable characteristics such as age, gender, and type of disease to patients' preference for shared decision making (e.g., Say, Murtagh, & Thomson, 2006). Nevertheless, the findings reported in this thesis (Chapters 1 and 2) show that a modifiable factor, that is, comprehension can promote the willingness and ability to participate in high-stakes value-sensitive decisions, independent of other influential factors (e.g., emotions and demographics).

Overall, the experimental results presented here echo previous calls for investing in citizens' and professionals' risk literacy, which is not only ethically desirable but can be also beneficial to public health goals (e.g., Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007). For example, participating in decision making about health has important benefits. Among others, risk literate citizens who can effectively participate in decision making demand fewer services, have higher treatment adherence, and improved confidence and coping skills (see Coultier & Collins, 2011).

One theoretical implication of these findings is that if we want to predict preference-sensitive decisions about health, traditional models about health behavior (e.g., the Health Belief Model (Rosenstock, 1974), the Theory of Reasoned Action (Fishbein, 1979)) should be updated to reflect the often central role of risk comprehension. For example, prominent models like the Social Cognitive Theory (Bandura, 2004) or the Information-Motivation-Behavioral Skills model (Fisher & Fisher, 2009) have addressed the role of health relevant knowledge in the prediction of health behavior. However, these models have not emphasized the role of numerical risk and benefit information that is common in decisions for which informed decision making is recommended (for an exception see Fuzzy Trace Theory, Reyna, Nelson, Han, & Dieckmann, 2009).

In addition, the majority of these models have been mostly based on and applied in the context of persuasion-based health promotion (i.e., encouraging a health behavior that is deemed desirable by experts). However, while some of the studied behaviors are still considered inherently positive, for other behaviors encouragement may no longer be ethically justified. Illustrative examples include screenings for breast and prostate cancer. In the past the norm was to encourage people to attend to these screening regularly. Hence, to address public health demands and experts' recommendations, most previous research on screening followed a persuasion-based tradition, looking to eliminate barriers to screening adherence (e.g., Austin, Ahmad, McNally, & Stewart, 2002; Curry & Emmons, 1994; Johnson, Mues, Mayne, & Kiblawi, 2008; Miller, Shoda, & Hurley, 1996). However, recently accumulated evidence shows that these screenings can cause substantial harms (e.g., Gøtzsche & Jørgensen, 2013; Ilic, Neuberger, Djulbegovic, & Dahm, 2013). These harms need to be weighed against potential benefits in the decision to get screened, making decisions about these cancer screenings preference-sensitive (Rimer et al., 2004). Moroever, essential ethical analyses in bioethics and the philosophy of informed decision making have long emphasized the importance of deliberation in the light of one's values precisely because there in no obvious "right" choice (Feltz, 2015). That is, many health decisions involve trade-offs that are highly subjective and can only be evaluated as "better" or "worse" relative to one's values (e.g., should you take a 20% chance of living 6 months longer with 50% lower quality of life)? In sum, while traditional persuasion-based models of health behavior may still apply to beneficial screenings with negligible harms (for example, cardiovascular risk screening in Chapter 8), to apply to procedures where informed decision making is recommended, they need to include comprehension of benefits and harms (Chapter 2).

Besides these theoretical implications, this thesis has identified several obstacles to informed decision making, as well as several strategies and tools that can facilitate comprehension and decision making. Various studies have shown that unfortunately not all patients and physicians have developed the skills that tend to be necessary for evaluating and understanding the complex statistics that are now part of many medical technologies and decisions (see Chapters 1, 3, and 4). We examined two types of abilities that can greatly contribute to risk literacy in this and other contexts, statistical numeracy and science literacy.

1.2. Numeracy helps both patients and physicians understand, feel, and evaluate trade-offs between benefits and risks.

Statistical numeracy refers to the ability to understand and evaluate numerical expressions of probabilities and risk, and has been related to superior decision making and risk literacy across many decision contexts, including health, education, and finance (e.g., Garcia-Retamero & Galesic, 2012; Cokely et al, 2012; Cokely & Kelley, 2009; Reyna, 2008). This thesis has demonstrated that numeracy promotes comprehension of complex and counter-intuitive evidence (Chapters 1 to 3), comprehension of risks from emerging high-profile threats (Chapter 4), and efficient decision making for high stakes decisions under time pressure (Chapter 6), independently of other important factors like education, demographics or emotions. Importantly, what these findings highlight is that informed judgments do not necessarily require complex calculations, abstract reasoning or exceptional levels of intelligence (Cokely et al., 2012; Ghazal, Cokely, & Garcia-Retamero, 2014): Instead, numeracy generally promotes good decision making in part because it is gives rise to more precise, adaptive emotional reactions (Chapter 5, Peters et al., 2006; Petrova, van der Pligt, & Garcia-Retamero, 2014), more intuitive understanding (Reyna et al., 2009), and simple yet deliberate metacognitive strategies that focus on the most essential information (Chapter IX; Cokely et al., 2012; Cokely & Kelley, 2009; Ghazal et al., 2014). For example, in general more numerate individuals tend to invest more time deliberating and thinking about thinking (and emotion) during decision making (Chapter 2, Garcia-Retamero, Cokely, Wicki, & Joeris, in press). Yet because this deliberate strategy tends to provide a better understanding of trade-offs, rewards, and risk, it also tends to be sensitive to concerns like time pressure and the possibility of a very bad outcome, leading to reliance on other simple yet adaptive strategies (Chapter 6).

1.3. Science literacy helps make sense of medical information and evidence.

Another ability that can contribute to comprehension of statistical evidence, independently of numeracy, is science literacy. Science literacy or scientific reasoning ability is the basic understanding of how science generates and assesses evidence (National Science Foundation, 2014). Chapter 2 demonstrates that science literacy can help individuals evaluate the evidence of benefits and harms of screening, and this effect is independent of that of numeracy. Ultimately, for medical decisions to be evidence-based, decision makers do not only need to calculate the risks and benefits, but also evaluate the strength of the evidence and give it the weight it deserves. This is especially important given that the public is aware that evidence for benefits and harms is not always and only provided by expert, well-intended sources (e.g., Chapter 10). Knowing the basics of how medical evidence is generated can also help the public distinguish low from high quality evidence and make sense of the information that is presented. For example, knowing the essence of experimental methods (e.g., a control group is required to establish the benefit of a treatment) can improve comprehension and adjust perceptions (Chapter 2). Furthermore, good science literacy may encourage people to approach rather than avoid health-relevant numerical information, and spend more time deliberating, which can in turn contribute to comprehension. Although the evidence for the benefit of science literacy is not as abundant as that for numeracy, the theory and results presented here suggest that it is a fruitful vein for future research and application in this domain (Drummond & Fischhhoff, 2015).

1.4. Visual aids are effective at increasing comprehension under certain conditions and are preferred by physicians.

Fortunately, solutions for deficiencies in numeracy and science literacy do not only lie in the future, and risk literacy is not only an individual trait characteristic. It is possible to inform and engage in decisions persons from all walks of life by providing well-designed, transparent information (e.g., Coultier & Collins, 2011; Garcia-Retamero & Galesic, 2013). One way to increase comprehension even among those who do not have sophisticated abilities is the provision of visual aids (Garcia-Retamero & Cokely, 2011; 2013; 2014). Difficulties in comprehension demonstrated in this thesis highlight that there is a pressing need for well-designed, transparent decision aids designed to help less numerate physicians and patients understand and discuss life-altering risks and benefits. For example, a simple visual aid representing the risk of contracting the Ebola virus and dying from it increased understanding, decreased fear, and improved judgment (Chapter 4). Even when the evidence that needs to be considered is more complex or counter-intuitive, user-friendly visual aids can be effective. For example, an icon array depicting the number of men who benefitted and were harmed from screening for prostate cancer with PSA tests helped men understand the trade-offs and make better decisions (Chapter 1). Chapter 3 also showed that physicians chose visual aids as the most preferred mode of communicating risk to low numeracy patients. This suggests that physicians are both trained and willing to use visual aids when communicating with patients. Even if they do not know that visual aids make risks easier to comprehend, these powerful and simple tools would be a ready risk communication means at their disposal.

While well-designed visual aids are in general an effective solution, there is no guarantee that they will be the best solution for everyone. For example, while icon arrays are among the most popular visual aids for risk communication, some types of icons are more effective than others (Zikmund-Fisher et al., 2014; but see also Gaissmaier et al., 2012) and interactive and visually appealing icon arrays can actually distract people and lead to worse outcomes (Zikmund-Fisher, Dickson, & Witteman, 2011). In addition, many people have severe diffiulties understanding visually represented information in graphs and charts (Okan, Garcia-Retamero, Galesic, & Cokely, 2012; Okan, Garcia-Retamero, Cokely, & Maldonado, 2015). Such difficulties can be potentially augmented under conditions of emotional distress, which tend to accompany important medical decisions. For instance, visual aids were not effective at increasing comprehension of benefits and harms from screening with mammography among women who perceived that breast cancer was an extremely severe diagnosis (Chapter 1). This result suggested that one's previous beliefs about the effectiveness of screening, emotions instilled by persuasive campaigns, or strong fears about diseases may interfere with shared and informed decision making, ideas that we pursued in Chapters 2 and 3 (see also Figure 1).

1.5. Beliefs and attitudes influence decisions about health, sometimes leading to non-evidence based decisions.

Much like anticipatory emotions, beliefs and attitudes can influence decisions about health (see Figure 1; see also Chapters 2 and 5). For example, many people can have strong positive beliefs about screening (Schwartz, Woloshin, Fowler, & Welch, 2004; Waller, Osborne, & Wardle, 2015). These beliefs could rightfully stem from the perceived value of saving a life or the high effectiveness of some screening programs. These beliefs can be further enforced by exposure to persuasive campaigns encouraging regular screening. As a result, people may be left with the impression that screening is useful by definition rather than a matter of choice based on a cost-benefit analysis. However, such attitudes can also lead to misinformed decisions about screening programs with disputed or mixed efficacy, where the benefits need to be carefully weighed against potential harms (e.g., Gøtzsche & Jørgensen, 2013; Ilic et al., 2013). Further, individuals may engage in motivated reasoning (Kunda, 1990) and trust evidence selectively in patterns that support their expectations; they can also discount or dismiss information that is contrary to their beliefs (Lewandowsky et al., 2012; Kahan, 2012). Consistent with these mechanisms, participants who had stronger positive a-priori beliefs about the goodness of screening, had stronger intentions to get screened, even when shown evidence that the screening was ineffective and potentially harmful (Chapter 2). Importantly, strong beliefs were not related to comprehension, suggesting that beliefs had no influence on how individuals processed the information; rather, the results were consistent with individuals giving more weight to the benefit of screening, consistent with their beliefs.

Attitudes and beliefs can also influence the decisions of surrogates (Chapter 5). The decisions that we make for others are strongly based on what we perceive to be the social norm (Stone & Allgaier, 2008) and we often use our own preferences as an anchor (Epley, Keysar, Van Boven, & Gilovich, 2004; Marks & Arkes, 2008). For example, given the same risk information, participants who themselves had positive attitudes towards risk taking in health made more conservative surrogate decisions compared to participants who had less positive attitudes towards risk taking. This suggests that risk takers may be aware of their frequently counter-normative attitudes towards risk, and hence they might have been motivated to make a decision for the other person that they perceived as more consistent with social norms. These results illustrate how our own beliefs or attitudes can influence not only the decisions we make for ourselves, but also those we make for other people.

The implications of these results are that well-intentioned decision surrogates can make unnecessarily risky or cautious decisions because of their beliefs or attitudes regarding the goodness of badness of certain behaviors. Similarly, beliefs stemming from well-intended persuasive campaigns could contribute via “transfer” to “non-evidence-based” decision making in other contexts. For example, campaigns promoting certain behaviors or treatments without specifying the extent of benefit or mentioning possible harms can instill in the public the belief that such behaviors are always and only beneficial. Once these beliefs exist, it may be difficult to change them: research shows that it is a real challenge to re-adjust people’s beliefs with new information (Lewandowsky et al., 2012). This could be especially difficult when the new information is counter-intuitive or surprising (e.g., harms form early detection). Going forward, we need to identify effective ways in which to communicate the changing evidence regarding certain medical procedures without losing the public’s trust.

1.6. Emotions can be both beneficial and detrimental to health decisions.

Making decisions about health based on evidence sounds like a cold calculating process. However, this is rarely the case, and especially so when the stakes are high – we are afraid of catching a deadly virus, we are worried about a potential cancer diagnosis, and we are terrified from having a heart attack. An ever-growing amount of research shows that such “immediate” or “anticipatory” emotions have profound effects on perceptions and decisions (Ferrer, Klein, Lerner, Reyna & Keltner, 2014, Finucane, Alkahami, Slovic, & Johnson, 2000; Loewenstein & Lerner, 2003; Loewenstein, Weber, Hsee, & Welch, 2001; Slovic, Finucane, Peters, & MacGregor, 2004; Slovic & Peters, 2006). Considering implications for informed decision making, this thesis shows the beneficial and detrimental effects of emotions on evidence-based decisions.

Interestingly, emotions can facilitate informed decision making through increasing motivation and “embodying” comprehension (i. e., they act as “spotlight” or “motivators”) (Peters, Lipkus, & Diefenbach, 2006). For example, participants who were more afraid of a disease, spent more time deliberating on the evidence, had better comprehension, and ultimately made better decisions (e.g., they did not intend to participate in a screening that was not effective, Chapter 2).

Emotions can be also beneficial in surrogate decision making. Empathy, our ability to take the perspective of others and imagine how they would feel regarding the communicated risks, can be especially beneficial when we make decisions for other people. For example, an approach to surrogate decision making called “advance directive” requires that surrogates follow the preferences stated by the person for whom they are deciding (Lawrence & Brauner, 2009). A similar approach, substituted judgment, requires surrogates to make a decision that the person would have made if they were able (Lawrence & Brauner, 2009). Thus, accurately predicting someone’s feelings towards risks and options and the ability to incorporate these into decisions are potentially essential in surrogate decision making (Hsee & Weber, 1997; Loewenstein, 2005). Participants who tended to show high empathic concern made decisions for others that were more consistent with what others wanted, relative to decisions for themselves (Chapter 5). Importantly, this was the case even when decision makers were encouraged by external forces to make a different decision than what the person wanted (i.e., the wishes of the person were contrary to accepted social norms). Hence, the capacity to empathize and feel what others would feel regarding different risks and options, is potentially crucial for accuracy of surrogate decisions, and demonstrates yet another way in which emotions can benefit decision making about health.

However, under extreme conditions emotions can be detrimental to informed decision making. For example, strong fear and worry about the Ebola virus was related to overly cautious behavioral intentions, like not going to work or keeping children away from school (Chapter 4). Evoking stronger negative emotional reactions towards a disease was related to more perceived benefit of screening for the disease, and stronger intentions to get screened for a cancer-like disease (Chapter 2). Most importantly, this effect persisted even when the screening was ineffective. This means that participants who were more afraid of the disease perceived benefit even from screening that did not reduce chances of dying, and were willing to expose themselves to unnecessary risks (Chapters 1 and 2). In this case, this effect of emotions could be considered detrimental and leading to inferior, "non-informed" decision making.

These results show that emotions are an integral part of decision making about health. They suggest that as risk communicators and educators we can use emotions in an ethically acceptable way to attract attention or increase interest in the information we want to communicate (Witte & Allen, 2000). However, we also need to keep in mind that finding the exact "healthy dose of fear" may be challenging, and strong emotions can also lead to overly cautious decisions (Chapters 1, 2, 4). Taking into account values and preferences, which are inevitable intertwined with emotions, is also an essential part of informed decision making. However, we should keep in mind that strong emotions may results in decisions that are comforting to the decision maker but are not the best practice according to common sense and experts (Chapters 1 to 4).

2. Section II: Field Research

The results of the field studies showed that numeracy, emotions, social support, and properly designed information about health can affect important health outcomes like seeking medical attention, as well as the presence or severity of disease.

In Chapter 6 we presented a field study of acute coronary syndrome (ACS) survivors. Results echoed the findings from our laboratory studies and showed that the benefits of numeracy extend beyond decisions that require comprehension of complex risks and benefits. Seeking medical attention on time in case of an ACS improves chances of survival and reduces the chances of complications that would require follow up care (e.g., Newby et al., 1996; Berger et al., 1999). However, many patients delay seeking medical attention thereby exposing themselves to higher risks of serious complications or death (Moser et al., 2006). Results showed that patients with higher numeracy were less susceptible to such misjudgments: they were quicker to call an ambulance or go to the hospital for ACS symptoms. These results suggest that patients with higher numeracy are more likely to make good decisions in the case of ACS and incur fewer costs to the health system. They also confirm theories that numeracy does not benefit judgment solely by facilitating complex computations, but rather by the adaptive selection of simple decision strategies that focus on the most essential information (Cokely et al., 2012; Cokely & Kelley, 2009; Ghazal et al., 2014).

Another field study with ACS patients reported in Chapter 7 extended our laboratory findings regarding the role of emotions in health. In contrast to the laboratory studies which looked mostly at anticipatory emotions regarding specific stimuli (e.g., worry about cancer), this research focused on the effects of stable emotional tendencies. In particular, we studied the relationship between the proneness to experience strong negative emotions but inhibit their expression, called Type D personality (Denollet, 1998), and cardiovascular disease severity. We identified a subgroup of patients at increased cardiovascular risk: patients who had Type D personality and a previous history of cardiovascular disease experienced a more severe ACS. Importantly, these patients also had a less healthy lipid profile, which at least partially accounted for the heightened severity of their disease. This suggests that their emotional tendencies may have prevented them from adhering to the behavior changes recommended in cardiovascular rehabilitation (e.g., healthy diet, more exercise, medication adherence, etc.). These results accord with other investigations showing that persons with such emotional propensity may be at higher risk of cardiovascular disease (e.g., O'Dell, Masters, Spielmans, & Maisto, 2011). These results also suggest that we are in need of interventions that can help these patients not fall behind, make good health decisions, and follow them through. For example, because of their gloomy nature Type D patients often lack social support (Polman et al., 2010; Williams et al., 2008), which is essential for health and may indirectly contribute to their failure to stay healthy. Providing social support to these individuals may be one way to increase success rates.

In Chapter 8 we report a nation-wide investigation that provides support for this hypothesis. We investigated the relationship between perceived social support and adherence to cardiovascular risk screening in representative samples of the Spanish population. Regular screening for cardiovascular risk is highly recommended because it can reduce cardiovascular events without incurring any substantial harm to the individuals who participate (Sheridan et al., 2003). We found that compared to individuals who lack social support, individuals who have social support have twice the odds of adhering to blood pressure and cholesterol screening. Importantly, the extent of this benefit was mostly independent of other common health determinants like social class, age, gender, cardiovascular history, and health behavior. People who lack social support are at higher risk of developing cardiovascular disease and dying from cardiovascular disease if they develop it (Barth et al., 2010). These results identify an important mechanism by which this effect is generated – adherence to primary and secondary prevention screening. These results suggest that provision of social support or increasing perceptions of social support can promote regular screening and thus decrease the risk of cardiovascular events. Until we know the exact mechanisms by which people without social support come to be at a disadvantage, we are in need of policies and interventions that encourage regular cardiovascular screening equitably among individuals with and without support. Examples would be policies and campaigns that do not rely on social ties for their success or require less initiative or action from the individual.

In Chapter 9 we reviewed 289 evidence-based interventions aiming to reduce sexually-transmitted infections (STI) and sexual risk behavior in adolescents in the United States. We conducted a meta-analysis of the interventions that assessed incidence of STIs at follow up and identified key characteristics of successful interventions. Results echoed those of our experimental studies in emphasizing the impact of communicating comprehensive, user-friendly, actionable information. Interventions teaching adolescents about the risks of sexual transmission reduced incidence of STIs, and even more so when this information was accompanied by training in the necessary skills to protect oneself. In particular, properly designed interventions teaching condom use skills and communication and negotiation skills can reduce STIs by 20 – 30%. In contrast, approaches that do not provide comprehensive information about the risks and means of protection, but instead discourage sexual behaviors altogether, are ineffective at reducing STIs (i.e., abstinence approaches, Trenholm et al., 2007). Although such interventions are not associated with worse health outcomes compared to control groups, the existence of an effective alternative suggests that they expose adolescents to unnecessary risks, and if implemented, should be combined with comprehensive education.

In Chapter 10 we presented findings from a cross-cultural qualitative study that explored young women's experiences with risk information and risk communication strategies regarding vaccination against the Human Papilloma Virus (HPV). Participants reported consulting numerous sources of health information and desired transparent, balanced information about potential benefits and risks. Most importantly, they approached information critically, questioning the motives of information providers, demonstrating that the trustworthiness of sources is an additional factor that influences how health information affects decisions. These experiences with risk communication demonstrated that participants were well aware of their responsibility to make good, rational decisions about health in face of different challenges like understanding complex medical information, balancing costs and benefits, and finding transparent, trustworthy information (see also Gray Brunton et al., 2014). In line with results reported previously (Chapters 1, 2, 4), these results suggest that well-designed and targetted information can reduce unwanted influences from other sources (e.g., emotions, information from ill-intended or non-expert sources).

3. Conclusion

The goal of this thesis was to contribute to the psychology of informed decision making, examining how cognitive, emotional, and social factors influence decisions about health and health outcomes. We identified groups at high risk and provided recommendations for designing cost-effective interventions that provide the necessary information, skills, and support to ensure risk literacy and improve health. The findings reported here can have important implications for health promotion and informed decision making. Low numeracy, strong negative emotions, misplaced beliefs, and lack of social support can indirectly lead to poor health outcomes. However, the delivery of evidence-based interventions (e.g., risk communication materials in the form of visual aids) can facilitate comprehension, even when the evidence is complex and coutner-intuitive and emotions run high. Comprehension can not only improve the public's judgments but also help individuals realize the importance of their own values and preferences for some high-stakes decisions. By facilitating informed and shared decision making about health, comprehension and risk literacy can increase patient satisfaction, reduce inefficiencies in the health system, and improve health outcomes.

4. References

Austin, L. T., Ahmad, F., McNally, M., & Stewart, D. E. (2002). Breast and cervical cancer screening in Hispanic women: A literature review using the health belief model. *Women's Health Issues, 12*(3), 122-128.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior*, *31*(2), 143-164.

Barth, J., Schneider, S., & von Känel, R. (2010). Lack of social support in the etiology and the prognosis of coronary heart disease: a systematic review and meta-analysis. *Psychosomatic medicine*, *72*(3), 229-238.

Berger, P. B., Ellis, S. G., Holmes, D. R., Granger, C. B., Criger, D. A., Betriu, A., ... & GUSTO-II Investigators. (1999). Relationship between delay in performing direct coronary angioplasty and early clinical outcome in patients with acute myocardial infarction results from the Global Use of Strategies to Open Occluded Arteries in acute coronary syndromes (GUSTO-IIb) Trial. *Circulation*, *100*(1), 14-20.

Betsch, C., Böhm, R., Airhihenbuwa, C. O., Butler, R., Chapman, G. B., Haase, N., ... & Nurm, Ü. K. (2015). Improving medical decision making and health promotion through culture-sensitive health communication: an agenda for science and practice. *Medical Decision Making*. Advance online publication. doi: 10.1177/0272989X15600434

Bodemer, N., Müller, S. M., Okan, Y., Garcia-Retamero, R., & Neumeyer-Gromen, A. (2012). Do the media provide transparent health information? A cross-cultural comparison of public information about the HPV vaccine. *Vaccine*, *30*(25), 3747-3756.

Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician–patient encounter: revisiting the shared treatment decision-making model. *Social Science & Medicine*, *49*(5), 651-661.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, *7*(1), 25-47.

Cokely, E.T., & Kelley, C.M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*, 20-33.

Coulter, A., & Collins, A. (2011). *Making shared decision-making a reality: no decision about me, without me.* King's Fund.

Curry, S. J., & Emmons, K. M. (1994). Theoretical models for predicting and improving compliance with breast cancer screening. *Annals of Behavioral Medicine*, *16*(4), 302-316.

Denollet, J. (1998). Personality and coronary heart disease: The type-D scale-16 (DS16). *Annals of Behavioral Medicine, 20*(3), 209-215.

DiClemente, R. J., Crosby, R. A., & Kegler, M. (Eds.). (2009). *Emerging theories in health promotion practice and research.* John Wiley & Sons.

Drummond, C., & Fischhoff, B. (2015). Development and Validation of the Scientific Reasoning Scale. *Journal of Behavioral Decision Making*. Advance online publication.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*(3), 327.

Esserman, L. J., Thompson, I. M., Jr,&Reid, B. (2013). Overdiagnosis and overtreatment in cancer: An opportunity for improvement. *JAMA, 310*(8), 797-798.

Feltz, A. (2015). Ethical Information Transparency and Sexually Transmitted Infections. *Current HIV Research*, *13*(5), 421-431.

Ferrer, R. A., Klein, W. M. P., Lerner, J. S., Reyna, V., & Keltner, D. (2014). Emotions and health decision making: Extending the Appraisal Tendency Framework to improve health and health care. *Behavioral Economics and Public Health*, 1-25.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1-17.

Fishbein, M. (1979). A theory of reasoned action: Some applications and implications. *Nebraska Symposium on Motivation*, *27*, 65-116.

Fisher, J. D., & Fisher, W. A. (2002). The information-motivation-behavioral skills model. In R. J. DiClemente, R.A. Crosby, & M.C. Kegler, *EmergingTheories in Health Promotion Practice and Research: Strategies for Improving Public Health* (pp. 40-70). San Francisco, CA: Jossey-Bass.

Fowler, F. J., Levin, C. A., & Sepucha, K. R. (2011). Informing and involving patients to improve the quality of medical decisions. *Health Affairs*, *30*(4), 699-706.

Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A. S., Broschinski, S., & Politi, M. C. (2012). Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information. *Health Psychology*, *31*(3), 286-296.

Garcia-Retamero, R., & Cokely, E. T. (2011). Effective communication of risks to young adults: using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied*, *17*, 270-287.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science*, *22*(5), 392-399.

Garcia-Retamero, R., & Cokely, E.T. (2014). The Influence of skills, message frame, and visual aids on Prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making, 27(2),* 179-189.

Garcia-Retamero, R., Cokely, E. T., Wicki, B., & Joeris, B. (in press). Improving surgeons' risk literacy with visual aids. *Patient Education and Counseling.*

Garcia-Retamero, R., & Galesic, M. (Eds.). (2013). *Transparent communication of health risks: Overcoming cultural differences*. Springer Science & Business Media.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*(2), 53-96.

Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *Cochrane Database of Systematic Reviews*,*6*, CD001877.

Gray Brunton, C., Farver, I., Jaeger, M., Lenneis, A., Parve, K., Patarcic, D., Petrova, D., Hogg, R., Kennedy, C., Garcia-Retamero, R., & Todorova, I. (2014). Young women's constructions of the HPV vaccine: a cross-cultural, qualitative study in Scotland, Spain, Serbia and Bulgaria. *International Journal of Behavioral Medicine*, *21*(1), 11-19.

Hsee, C. K., & Weber, E. U. (1997). A fundamental prediction error: Self–others discrepancies in risk preference. *Journal of Experimental Psychology: General*, *126*(1), 45-53.

Ilic, D., Neuberger, M. M., Djulbegovic, M., & Dahm, P. (2013). Screening for prostate cancer. *Cochrane Database of Systematic Reviews,1*, CD004720.

Johnson, C. E., Mues, K. E., Mayne, S. L., & Kiblawi, A. N. (2008). Cervical cancer screening among immigrants and ethnic minorities: A systematic review using the Health Belief Model. *Journal of Lower Genital Tract Disease, 12*(3), 232-241.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin,108*(3), 480-498.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732-735.

Lawrence, R. E., & Brauner, D. J. (2009). Deciding for others: limitations of advance directives, substituted judgment, and best interest. *Virtual Mentor*, *11*(8), 571.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106-131.

Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. *Health Psychology*, *24*(4S), S49-S56.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. Davidson, K. Scherer & H. Goldsmith (Eds.), *Handbook of affective science* (pp. 619-642). New York: Oxford University Press.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267.

Marks, M. A., & Arkes, H. R. (2008). Patient and surrogate disagreement in end-of-life decisions: Can surrogates accurately predict patients' preferences? *Medical Decision Making, 28*(4), 524-531.

Miller, S. M., Shoda, Y., & Hurley, K. (1996). Applying cognitive-social theory to health-protective behavior: Breast self-examination in cancer screening. *Psychological Bulletin, 119*(1), 70-94.

Moser, D. K., Kimble, L. P., Alberts, M. J., Alonzo, A., Croft, J. B., Dracup, K., ... & Mensah, G. A. (2006). Reducing delay in seeking treatment by patients with acute coronary syndrome and stroke a scientific statement from the American heart association council on cardiovascular nursing and stroke council. *Circulation*, *114*(2), 168-182.

Moynihan, R., Bero, L., Ross-Degnan, D., Henry, D., Lee, K., Watkins, J., ... & Soumerai, S. B. (2000). Coverage by the news media of the benefits and risks of medications. *New England Journal of Medicine*, *342*(22), 1645-1650.

National Science Foundation. (2014). *Science and engineering indicators.* Washington, D.C.

Newby, L. K., Rutsch, W. R., Califf, R. M., Simoons, M. L., Aylward, P. E., Armstrong, P. W., ... & Van de Werf, F. (1996). Time from symptom onset to treatment and outcomes after thrombolytic therapy. *Journal of the American College of Cardiology*, *27*(7), 1646-1655.

O'Dell, K. R., Masters, K. S., Spielmans, G. I., & Maisto, S. A. (2011). Does type-D personality predict outcomes among patients with cardiovascular disease? A meta-analytic review. *Journal of Psychosomatic Research 71*(4), 199-206.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2015). Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *Journal of Experimental Psychology: Applied*, *21*(2),178-194

Okan, Y., Garcia-Retamero, R., Galesic, M., & Cokely, E. T. (2012). When higher bars are not larger quantities: On individual differences in the use of spatial information in graph comprehension. *Spatial Cognition & Computation*, *12*(2-3), 195-218.

Peters, E., Lipkus, I., & Diefenbach, M. A. (2006). The functions of affect in health communications and in the construction of health preferences. *Journal of Communication*, *56(s1)*, S140-S162.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*(5), 407-413.

Petrova, D., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making, 27*, 191-199.

Polman, R., Borkoles, E., & Nicholls, A. R. (2010). Type D personality, stress, and symptoms of burnout: The influence of avoidance coping and social support. *British Journal of Health Psychology*, *15*(3), 681-696.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making.*Psychological bulletin*, *135*(6), 943.

Rimer, B. K., Briss, P. A., Zeller, P. K., Chan, E. C., & Woolf, S. H. (2004). Informed decision making: what is its role in cancer screening? *Cancer*, *101*(S5), 1214-1228.

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education & Behavior*, *2*(4), 328-335.

Salzburg Global Seminar. (2011). The Salzburg Statement on Shared Decision Making. Retrieved on 03.02.2016 from https://www.lumc.nl/sub/2050/att/1405151019245212.pdf

Say, R., Murtagh, M., & Thomson, R. (2006). Patients' preference for involvement in medical decision making: a narrative review. *Patient Education and Counseling*, *60*(2), 102-114.

Sheridan, S., Pignone, M., & Donahue, K. (2003). Screening for high blood pressure: a review of the evidence for the US Preventive Services Task Force. *American Journal of Preventive Medicine*, *25*(2), 151-158.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, *24*(2), 311-322.

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*(6), 322-325.

Stokols, D. (1992). Establishing and maintaining healthy environments: toward a social ecology of health promotion. *American Psychologist*, *47*(1), 6-22.

Stone, E. R., & Allgaier, L. (2008). A social values analysis of self–other differences in decision making involving risk. *Basic and Applied Social Psychology, 30*(2), 114-129.

Schwartz, L. M., Woloshin, S., Fowler, F. J., & Welch, H. (2004). Enthusiasm for cancer screening in the United States. *JAMA, 291*(1), 71-78.

Trenholm, C., Devaney, B., Fortson, K., Quay, K., Wheeler, J., & Clark, M. (2007). *Impacts of Four Title V, Section 510 Abstinence Education Programs. Final Report*. Mathematica Policy Research, Inc.

Waller, J., Osborne, K., & Wardle, J. (2015). Enthusiasm for cancer screening in Great Britain: a general population survey. *British Journal of Cancer*, *112*(3), 562-566.

Williams, L., O'Connor, R. C., Howard, S., Hughes, B. M., Johnston, D. W., Hay, J. L., et al. (2008). Type-D personality mechanisms of effect: The role of health-related behavior and social support. *Journal of Psychosomatic Research, 64*(1), 63-69.

Wills, C. E., & Holmes-Rovner, M. (2003). Patient comprehension of information for shared treatment decision making: state of the art and future directions. *Patient Education and Counseling*, *50*(3), 285-290.

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education & Behavior*, *27*(5), 591-615.

Zikmund-Fisher, B. J., Dickson, M., & Witteman, H. O. (2011). Cool but counterproductive: interactive, web-based risk communications can backfire. *Journal of Medical Internet Research*, *13*(3), e60.

Zikmund-Fisher, B. J., Witteman, H. O., Dickson, M., Fuhrel-Forbis, A., Kahn, V. C., Exe, N. L., ... & Fagerlin, A. (2014). Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Medical Decision Making*, *34*(4), 443-453.

Curriculum Vitae

Dafina Petrova was born in the small town of Razgrad, Bulgaria. In 2007 she graduated from the public English Language High School Geo Milev in Rousse, Bulgaria. In 2010 she graduated from the BA program in Integrated Social and Cognitive Psychology at Jacobs University in Bremen, Germany. She then moved to the Netherlands to complete a research master in Psychology at the University of Amsterdam. In 2012 Dafina started her predoctoral fellowship at the University of Granada. For her thesis Dafina spent three months at Michigan Technological University in the USA and four months at Imperial College London in the United Kingdom. Dafina is a member of the Learning, Emotion, and Decisions Research Groups at the University of Granada and the Decision Psychology Lab at the University of Oklahoma.

Acknowledgments

I would like to first and foremost wholeheartedly thank my thesis advisers Rocio Garcia-Retamero, Andrés Catena, and Edward Cokely for teaching me, guiding me, and supporting me in the small and big decisions; for staying always positive, encouraging, and constructive. It is a pleasure and a privilege to learn from you, and an honor to be your student. I want to especially thank Rocio for giving the opportunity to complete this work and being supportive all the way.

I would like to thank all my co-authors for their invaluable contributions. Especially to Joop van der Pligt, who is unfortunately no longer among us. Joop was a great mentor and person, and was thus greatly missed while we were completing this work. My thanks go to all members – professors and students – of the Learning, Emotion, and Decisions Research Group at the University of Granada. They listened to my numerous presentations, provided valuable feedback, and supported me. Thanks to my friends and colleagues at the CIMCYC among whom Yasmina, David, Alberto, Paola, Manu and Raul, who were the closest thing I had to a family so far away from home. More thanks go to the members of Edward´s lab at Michigan Tech, to Olga Kostopoulou and Brendan Delaney at Imperial College London, to members of the Junior Researcher Programme, and to our collaborators at the Hospital Virgen de las Nieves. Special thanks to all cardiac patients who volunteered to participate in our research during a very difficult time in their lives and for no remuneration. I am indebted to the members of the examination committee and the external reviewers for reviewing this thesis. I also would like to thank the anonymous reviewers of the published papers for their constructive comments. Many of the included papers were greatly improved by the anonymous feedback.

Finally, thanks to my mom, for being my best friend and for making me believe I can do anything.