UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

NUMERATE PEOPLE BENEFIT MORE FROM REGULAR SCIENCE REPORTING: THE CRITICAL ROLE OF SCIENTIFIC REASONING AND CAUSAL MISUNDERSTANDING

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

OLIVIA D. PERRIN
Norman, Oklahoma
2024

NUMERATE PEOPLE BENEFIT MORE FROM REGULAR SCIENCE REPORTING: THE CRITICAL ROLE OF SCIENTIFIC REASONING AND CAUSAL MISUNDERSTANDING

A THESIS APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Edward T. Cokely, Chair

Dr. Adam Feltz

Dr. Rocio Garcia-Retamero

Acknowledgments

This work would not have been possible without the support and guidance of my advisor, Dr. Edward T. Cokely. I would like to thank my committee members, Dr. Adam Feltz and Dr. Rocio Garcia-Retamero, for offering their time, advice, and feedback on this project. I would also like to extend my gratitude to my friends and lab mates at the University of Oklahoma for providing support and encouragement throughout the process. Finally, I would like to extend a special thanks to my partner and family for their continuous love and support. To my brother Nick, thank you for always inspiring me to continue growing and learning.

Table of Contents

List of Tables

List of Figures

Abstract

Numerate people tend to make more informed judgments and decisions because they are more risk literate (i.e., better able to evaluate and understand risk). Do numeracy skills also help people understand regular science reporting from mainstream news sources? To address this question, we investigated responses to regular science reports (e.g., excerpts from CNN health), testing a cognitive model linking numeracy, scientific reasoning, judgment biases, and casual theory errors (i.e., misinterpreting correlational information as causal; Seifert et al., 2022). In Study 1 (n=200), structural equation modeling indicated that numerate people were less likely to exhibit judgment biases because they were better at scientific reasoning, which helped them avoid causal misinterpretations. Study 2 (n=342) cross-validated findings from Study 1, indicating that the link between numeracy and scientific reasoning was also associated with improved cognitive self-assessment (e.g., reduced overconfidence on comprehension judgments). Results suggest numerate people may generally be less likely to confuse correlation and causation in regular science reporting. Results also suggest that numerate people are more likely to have acquired scientific reasoning skills that more generally support risk literacy and knowledge acquisition, consistent with the Knowledge is Power account of Skilled Decision Theory (Cokely et al., 2018). Discussion focuses on implications for risk literacy theory and training, and includes a Risk Literacy Difficulty Analysis indicating that nearly half of the US adult population may be likely to misunderstand common types of regular science reports.

Keywords: Numeracy, risk literacy, scientific reasoning, causal theory errors

Chapter 1: Introduction

Many people learn about scientific findings from regular science reporting—i.e., typical journalistic reports created and distributed by mainstream news and mass media providers (Funk et al., 2017; Hendriks et al., 2020; National Science Board, 2018; Pavelle & Wilkinson, 2020). However, research indicates people's ability to accurately interpret regular science reporting varies as a function of individual differences in education, knowledge, skills, abilities, and cognitive styles (e.g., Bromme & Goldman, 2014; Miller, 1998; Stanovich & West, 1997). For example, individual differences in specialized scientific reasoning skills appear to influence people's ability to interpret and understand the quality of scientific evidence (Drummond & Fischhoff, 2017). As such, some people may be much more likely to misunderstand some scientific information about personally relevant risks, which in turn could lead to costly judgment errors and biases (e.g., misinformed attitudes about risk mitigation).

Theoretically, statistical numeracy skills (i.e., practical probabilistic and inductive reasoning skills) may be another factor that could help people avoid misinterpretations of regular science reporting because numeracy is one of the most robust predictors of general decision making skill including the ability to evaluate and understand risk (i.e., risk literacy; Cokely et al., 2012; 2018; see RiskLiteracy.org). There are many cognitive mechanisms thought to help explain the link between statistical numeracy and decision quality (e.g., Cokely & Kelley, 2009; Garcia-Retamero et al., 2019; Peters, 2012; 2020; Peters et al., 2006; Petrova et al., 2023; Reyna et al., 2009). However, according to the Knowledge is Power account of Skilled Decision Theory (Cokely et al., 2018; 2025; see also Cho et al. 2023), variation in knowledge (i.e., *representative understanding*) is by far the most influential factor that typically gives rise to differences in skilled judgment and decision making among experts and non-experts alike (e.g., Chase & Simon, 1973;

Ericsson et al., 2006; 2007; 2018; Klein, 1999). Put more simply, because numerate people are more risk literate, they tend to acquire more accurate knowledge about many risk-relevant topics, which in turn promotes more informed and therefore less biased judgment and decision making.

Given that numerate people are better able to independently acquire, evaluate, and understand risk information, one open question is whether numerate people might also benefit more from regular science reporting. Some evidence suggests this kind of relationship may be likely. Numerate people tend to correctly interpret many kinds of professionally developed risk and science communications (e.g., health, finance, and safety; Garcia-Retamero & Cokely, 2013; 2017; Garcia-Retamero et al., 2020; National Academies of Sciences, 2017; Peters, 2012; 2020; Reyna et al., 2009; see also Bruine de Bruin & Bostrom, 2013). Numerate people are less susceptible to biases related to how numerical information is presented (Garcia-Retamero & Cokely, 2011; 2014; 2015; Garcia-Retamero & Galesic, 2010; Peters & Bunquin, 2024; Petrova et al., 2015; 2016) and are less susceptible to misinformation and motivated reasoning effects associated with the use of some alternative or social media sources (Cho et al., 2023; Kahne & Bowyer, 2017; Roh et al., 2015; Roozenbeek et al., 2020; Taber & Lodge, 2006; Van der Linden, 2023). Numerate people also tend to be less susceptible to folk science myths and generally acquire more accurate knowledge about many diverse hazards and risks (Allan et al., 2017; Feltz, 2016; Feltz et al. 2022; 2023; Garrido et al., 2021; Ghazal et al. 2014; Grounds & Joslyn, 2018; Michal & Shah, 2024; Okan et al. 2012; 2016; Petrova et al. 2017; 2018; Ramasubramanian et al., 2019; Tanner & Feltz, 2022; Tanner et al., 2024). However, currently there do not appear to be any direct tests of the relations among numeracy, scientific reasoning, judgment biases, and comprehension of regular science reporting.

Is Numeracy Related to Scientific Reasoning?

Parallels between numeracy and scientific thinking have been discussed for more than 60 years. In the earliest known formal conceptualization of numeracy, the Crowther Report discussed numeracy as a "mirror image of literacy" and proposed understanding the scientific method as a possible component of numeracy (Ministry of Education, 1959, p. 270). More recent reports have also characterized numeracy skills as foundational prerequisites for being scientifically literate (National Academies of Sciences, 2016). Furthermore, some studies have noted that one must be able to use and apply mathematical knowledge to be able to engage in scientific reasoning and related problem-solving activities (e.g., proportional reasoning; Hilton & Hilton, 2016). However, specific theories about the nature of scientific reasoning have also been a topic of independent psychological inquiry for many years, including as part of Inhelder and Piaget's (1958) theory of cognitive development. While that early theory assumed scientific reasoning was a single cognitive ability akin to a general latent trait with broad applications, later conceptualizations favored more multidimensional constructs (Optiz et al., 2017). For example, influenced by Newell and Simon's (1972) theory of problem solving, Klahr and Dunbar (1988) developed an integrated model of scientific reasoning processes (*Scientific Discovery as Dual Search – SDDS*), which emphasized both domain-specific knowledge and general cognitive mechanisms. According to Klahr and Dunbar (1988), scientific reasoning primarily involved guided search within two related, but distinguishable problem spaces – a *hypothesis space* and an *experiment space*.

More recently, scientific reasoning has been viewed as a concept that may contain multiple independent characteristics or skills (e.g., the eight skill model from Fischer et al., 2014; but see Optiz et al., 2017 for issues related to debates about the role of domain general v. domain specific skills). Several modern theories emphasize independent roles of distinct general-purpose cognitive operations including induction, deduction, analogy, problem solving, and causal reasoning (Barz & Achimaş-Cadariu, 2016; Dunbar & Fugelsang, 2005; Kambeyo & Csapó, 2018). This view is generally consistent with broader modern conceptualizations of scientific reasoning defined primarily as the thinking and reasoning skills involved in (1) scientific inquiry, (2) experimentation, (3) evidence evaluation, (4) inference, and (5) argumentation (Zimmerman, 2000; but for related notions of scientific thinking and scientific discovery see also Bao et al., 2022; Díaz et al., 2023).

More than a dozen scientific reasoning tests have been developed since the turn of the century using a variety of test formats (e.g., closed v. open ended; Optiz et al., 2017). Most of the modern tests are primarily intended to assess individual differences in skill achievement levels among high school and university students. However, the Scientific Reasoning Scale (SRS; Drummond & Fischhoff, 2017) is different because it was specifically designed and validated for broader use with educated adult samples (e.g., adults in industrialized countries). Recent studies have also directly examined the association between numeracy and scientific reasoning.[1] For example, in a series of four studies, Drummond & Fischhoff (2017) found that numeracy was significantly correlated with scores on the SRS (r = .28, .36, and .40; Drummond & Fischhoff, 2017). Additionally, Drummond & Fischhoff (2019) conducted a study to assess whether priming scientific reasoning skills might influence a person's myside bias when evaluating scientific evidence and found that numeracy was significantly related with SRS scores (r = .47). In another study that examined factors that might inhibit scientific reasoning skills (e.g., the content of the problems), numeracy was again found to be significantly correlated with SRS scores (r = .42;

[1] Other related concepts such as science literacy have also been shown to be related to numeracy, in the context of complex medical screening decisions (Petrova et al., 2016; 2019) and perceptions of immigration (Savadori et al., 2016).

Bašnáková et al., 2021). Taken together, numeracy appears to be robustly and systematically related to scientific reasoning and related concepts. However, much less is known about the cognitive mechanisms that link numeracy and scientific reasoning with regular science reporting and potential downstream judgment biases.

Why Do People Misunderstand Regular Science Reporting?

Recent research suggests that one common source of misunderstanding in regular science reporting may be *causal theory errors* (i.e., inferring or accepting a causal claim based on correlational evidence alone; Seifert et al., 2022). For example, results from Seifert et al. (2022) indicated that 63 percent of their participants erroneously accepted a causal claim based on correlational evidence alone. This finding accords with others that have shown similarly high rates of endorsing causal claims based on correlational evidence (Bleske-Rechek et al., 2015; Xiong et al., 2019; see also Adams et al., 2017; Norris et al., 2003). Given that science stories in the media have been linked to changes in both health-related behavior and knowledge (e.g., cancer screening decisions; Cram et al., 2003; Matthews et al., 2016; Stryker et al., 2008), these results seem to suggest that causal theory errors might lead to misunderstandings that could result in costly downstream attitude and judgment biases.

It is currently unclear whether numeracy or scientific reasoning skills are related to individual differences in susceptibility to causal theory errors, although it appears theoretically likely. However, recent research suggests that interventions can reduce causal theory errors. For example, some studies have focused on the actions of science writers (e.g., journalists, researchers), suggesting that improving data visualization designs (Xiong et al., 2019) or including caveats (Bott et al., 2019) may help mitigate unwarranted perceptions of causality (see also Bratton et al., 2019; 2020; Sumner et al., 2014; 2016; Woloshin et al., 2009). Alternatively, other studies have focused on interventions that directly aim to improve people's reasoning when evaluating science studies (e.g., encouraging consideration of alternative causal theories; Seifert et al., 2022), which significantly reduced the occurrence of causal theory errors on regular science reporting type tasks. The benefits of the reasoning intervention suggest that differences in cognitive strategies or knowledge can create differences in understanding and judgment. Furthermore, there is some distant yet conceptually related evidence suggesting that differences in causal knowledge can simplify and inform related decision making processes (Garcia-Retamero & Hoffrage, 2006; Garcia-Retamero et al., 2007). Theory also suggests that scientific reasoning and causal reasoning should generally be related, which is consistent with the notion that both numeracy and scientific reasoning may be important skills that could reduce causal theory errors (Kuhn, 2012; Kuhn & Dean, 2004).

Despite potential connections, it remains unknown whether statistical numeracy and scientific reasoning skills are reliably related to causal understanding of regular science reporting. Given that numeracy promotes people's ability to independently acquire knowledge and become more informed decision makers, I hypothesize that numerate people may also have acquired other skills (e.g., scientific reasoning skills), in order to support their ability to acquire relevant information and make more informed judgments and decisions more generally (i.e., the *Risk Literacy Skills for Knowledge Hypothesis*). In other words, I predict that numerate people may avoid causal theory errors and downstream judgment biases primarily because they have acquired specialized scientific reasoning skills, which is theoretically consistent with the Knowledge is Power account of Skilled Decision Theory (Cokely et al., 2018; 2025).

Current Studies

The primary aim of the current studies was to provide the first direct test of a cognitive model based on Skilled Decision Theory, mapping the interrelations among numeracy, scientific reasoning, judgment biases and causal theory errors in response to regular science reporting. Specifically, Study 1 presented participants with four real regular science report excerpts (e.g., an article from CNN health), and tested a structural equation model of the relations between the variables of interest. Study 2 then replicated the model from Study 1, providing an out-of-sample cross-validation test, and extending the investigation to test the impact of participants' metacognitive self-assessments on downstream judgment biases (i.e., susceptibility to overconfidence in science report comprehension).

Chapter 2: Study 1

Participants and Procedure

The data were collected via Qualtrics in Spring 2023 using an undergraduate student sample. The study was designed to take roughly one hour to complete, and the average completion time was 40 minutes. Out of 276 total participants, 200 participants were retained for the analyses after excluding participants who took insufficient time (i.e., those who took less than 20 minutes to complete the study). Of the 200 participants, 160 (80%) identified as female, 32 (16%) identified as male, 2 (1%) identified as non-binary, and 6 (3%) selected “prefer not to say.” Participants reported ages ranged between 18 and 33 years old (M = 19.4, SD = 1.8).

Participants read four news article excerpts selected from online new sources (e.g., CNN Health), which were presented in a randomized order (Cosdon, 2022; Elliot, 2022; Kissell, 2022; Woodyatt, 2021; see Appendix for excerpts).[2] Each excerpt reported on scientific findings that had recently been published in peer-reviewed journals. The articles were selected because they each presented a correlational relationship between two variables that were apolitical in nature and discussed implications for health or other consequences (e.g., increased risk of dementia from watching too much TV). As such, each article presented an opportunity for a causal theory error. After each excerpt, participants answered questions related to causal misunderstanding (i.e., the extent to which they agreed correlational information presented in the article was causal), reported risk perceptions (i.e., how concerned they think someone should be about the risk), and made judgments about multiple behavioral intentions (see below).

[2] Study 1 included a between subjects experimental manipulation, which was ultimately not significant. Therefore, all following analyses are collapsed across conditions.

Measures

Statistical Numeracy

The Berlin Numeracy Test (BNT; Cokely et al., 2012) was used to assess numeracy and risk literacy. Following best-practice recommendations, this study included the BNT-S form, which includes four items from the BNT and three items from Schwartz et al. (1997) and provides increased sensitivity by allowing for a wider range of skill assessment. The sum of correct answers (0 – 7) was used as the statistical numeracy score.

Scientific Reasoning

Scientific reasoning was assessed by the 11-item Scientific Reasoning Scale (Drummond & Fischhoff, 2017), which is an individual difference measure of the skills needed to evaluate the quality of scientific findings. Each item is comprised of a short scientific scenario followed by a statement that respondents must evaluate as either true or false. The sum of correct answers (0 – 11) was used as the scientific reasoning score.

Causal Misunderstanding

Following Seifert and colleagues (2022), causal theory errors were operationalized as causal misunderstanding. Specifically, I measured participant's responses to one item asking the extent to which participants agreed that the information presented in the article was causal (e.g., *According to the article, watching a lot of TV causes dementia*).[3] The scale ranged from 1 (Not at all) to 6 (Strongly Agree). The mean rating of this statement across four article excerpts was used as the score for causal misunderstanding, with higher scores indicating more causal misunderstanding.[4]

[3] Because it is possible that people may have prior relevant knowledge, we explicitly directed participants to respond solely based on the information provided in the article excerpts.
[4] Two items were originally used to assess causal misunderstanding in Study 1. However, one item was dropped as the wording seemed to elicit confusion.

Risk Perceptions

Risk perceptions were measured with one item asking about the extent to which people should be concerned about the risk (e.g., *How concerned should someone who watches a lot of TV be about getting dementia?*). The scale ranged from 0 (Not at all concerned) to 6 (Extremely concerned). The mean rating of this item across four article excerpts was used as the score for risk perceptions, with higher scores indicating more risk perception bias (i.e., expressing more concern).

Judgment Biases

Judgment biases were modeled as a latent trait using four items that assessed various possible judgments about behavioral intentions after reading each article excerpt (e.g., conditional intention to change one's own behavior). The scale for each item ranged from 1 (Not at all) to 6 (Strongly Agree), with higher scores indicating more biased judgment. The four items are described in detail below:

Personal. This item assessed the extent to which participants agreed that they would change their own behavior based on information presented in the article. An example item is, "*Given the information from the article, I would reduce the amount of TV that I watch to prevent dementia.*" The mean rating of this statement across four article excerpts was used as the score for "personal."

Recommendations. This item assessed the extent to which participants agreed that they would recommend a change in behavior to their friends and family. An example item is, "*Given the information from the article, I would recommend that my friends and family reduce the amount of TV that they watch.*" The mean rating of this statement across four article excerpts was used as the score for "recommendations."

Social Media. This item assessed the extent to which participants agreed that they would share information about the risk on social media. An example item is, "*Given the information from the article, I would share on social media that it is important to reduce the amount of TV watched in order to prevent dementia*." The mean rating of this statement across four article excerpts was used as the score for "social media."

Policy. This item assessed the extent to which participants agreed that they would redirect funding within a local community. An example item is, "*A local community has a risk communication program that aims to improve people's health. They are considering redirecting some funding for HIV/AIDS risk communication toward risk communication that gives a warning about the relationship between watching TV and dementia. Based on the information from the article, they should choose to redirect the funding.*" The mean rating of this statement across four article excerpts was used as the score for "policy."

Results and Analyses

All statistical analyses were completed in R version 4.3.0. Base R packages were used to compute descriptive statistics and correlations for all Study 1 variables (Table 1). To test relationships between numeracy, scientific reasoning, causal misunderstanding, and downstream consequences, a structural modeling approach was used, and a latent factor of judgment biases was estimated and included in the structural model (Figure 1). Furthermore, relevant indirect effects within the model were tested and presented with 95% Confidence Intervals (CI) estimated using 1,000 bootstrap samples (Table 2).

Table 1

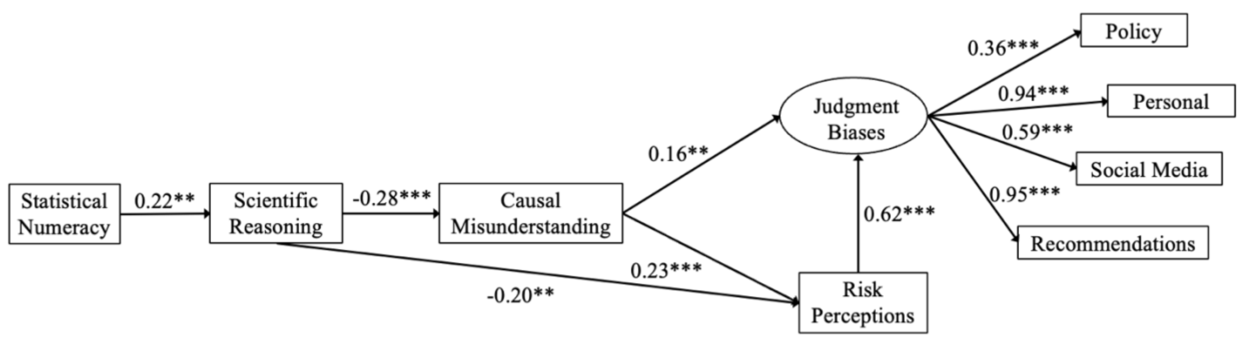
Descriptive Statistics and Correlations for Study 1 Variables

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. BNT-S	3.14	1.67								
2. SRS	5.68	2.45	.22**							
3. Causal Misunderstanding	3.59	1.06	-.08	-.28**						
4. Risk Perceptions	3.40	0.85	.01	-.27**	.29**					
5. Judgment Biases	3.42	0.74	.02	-.25**	.38**	.66**				
6. Social Media	2.80	1.11	.01	-.25**	.29**	.51**	.76**			
7. Recommendations	3.87	0.85	.05	-.19**	.32**	.63**	.87**	.55**		
8. Personal	3.83	0.85	.02	-.23**	.30**	.61**	.86**	.56**	.89**	
9. Policy	3.19	1.05	-.02	-.10	.27**	.32**	.62**	.20**	.35**	.30**

Note. *$p < .05$, **$p < .01$

Figure 1

Structural Equation Model with Causal Misunderstanding



Note. Reported are standardized coefficients. *$p < .05$, **$p < .01$, ***$p < .001$. Model estimated using the *lavaan* package in R (Rosseel, 2012).

As shown in Figure 1, structural equation modeling suggested numerate people were less likely to misunderstand correlational information as causal specifically because they tended to be better at scientific reasoning. In accord with accepted standards for model evaluation (Hu & Bentler, 1999), this model demonstrated good fit, $\chi^2_{18} = 29.17$, $p = 0.046$, CFI = 0.98, TLI = 0.97, RMSEA = 0.06 [0.01, 0.09], SRMR = 0.05. Specifically, statistical numeracy significantly predicted scientific reasoning ($\beta = 0.22$), which in turn significantly predicted causal misunderstanding (β = -0.28). Causal misunderstanding also partially mediated the relationship between scientific reasoning and risk perceptions. Furthermore, causal misunderstanding partially mediated the relationship between scientific reasoning and judgment biases (see Figure 1).

As shown in Table 2, an analysis of indirect effects indicated that numeracy exerted a significant indirect effect on causal misunderstanding (-.06, 95% CI [-.12, -.02]) as well as on risk perceptions (-.06, 95% CI [-.11, -.02]). Finally, numeracy exerted a significant indirect effect on judgment biases, through scientific reasoning, causal misunderstanding, and risk perceptions (-.05, 95% CI [-.09, -.02]).

Table 2

Total Indirect Effects (Standardized Coefficients) for Key Variables in Study 1

Paths	Estimate	SE	Bootstrapped 95% CI
(1) Numeracy → scientific reasoning → **causal misunderstanding**	-.06*	.03	[-.12, -.02]
(1) Numeracy → scientific reasoning → **risk perceptions** (2) Numeracy → scientific reasoning → causal misunderstanding → **risk perceptions**	-.06*	.02	[-.11, -.02]
(1) Numeracy → scientific reasoning → risk perceptions → **judgment biases** (2) Numeracy → scientific reasoning → causal misunderstanding → **judgment biases** (3) Numeracy → scientific reasoning → causal misunderstanding → risk perceptions → **judgment biases**	-.05**	.02	[-.09, -.02]

Note. Estimate reflects the total indirect effect from multiple paths for each dependent variable. Effects were estimated with 1,000 bootstrap samples. SE = standard error; CI = confidence interval. **p* < .05, ***p* < .01.

Study 1 Discussion

Results from Study 1 indicate that numerate people were indeed more likely to benefit from regular science reporting as compared to less numerate people. Specifically, SEM results suggested that more numerate people were less likely to misconstrue correlational information as causal, which in turn led to fewer downstream biases (i.e., less biased risk perceptions and judgments).

Critically, these observed benefits of numeracy followed because numerate people had previously acquired specialized scientific reasoning skills. Theoretically, this evidence suggests that one reason numerate people may become more knowledgeable and informed decision makers is that they may be more likely to acquire some additional risk literacy skills (i.e., scientific reasoning) that sometimes help facilitate risk understanding and knowledge acquisition.

While these findings were largely consistent with anticipated results, I had specific concerns about two primary limitations of the current study. First, results suggested that some participants may not have complied with instructions (e.g., 67% of participants failed one of two attention checks). This potentially careless responding could have implications for the reliability and interpretability of the observed model fit (see Voss, 2023 for a brief review), which suggests that replication and cross-validation is merited. Second, it is possible that using a strength of agreement Likert scale to assess causal misunderstanding may have weakened the psychometric precision of the variable.[5] Although using a Likert scale has some advantages (e.g., wider range of response options), measuring causal misunderstanding in this way made it more difficult to precisely or unequivocally estimate causal theory errors. Furthermore, given the influential role of knowledge outlined in other recent research (e.g., Cho et al., 2023), it seems likely that, provided sufficient psychometric sensitivity, causal misunderstanding should fully mediate the relationship between scientific reasoning and downstream consequences (i.e., risk perceptions and judgment biases).

[5] See the general discussion for a detailed discussion of another potential limitation, namely the non-significant direct effect of numeracy, which also appears to have been the result of a psychometric measurement limitation (see also Study 2).

Chapter 3: Study 2

To begin to address potential measurement limitations of Study 1, Study 2 included an updated measure of causal misunderstanding and an additional test of a metacognition bias (i.e., overconfidence), which was investigated using two of the four article excerpts from Study 1. Specifically, in Study 2 causal misunderstanding was measured via a single binary choice item, allowing for a more definitive measure of causal theory errors. In addition, Study 2 aimed to address previous research suggesting that metacognitive biases may affect people's ability to acquire new information because they may limit people's ability to evaluate their own comprehension (e.g., overconfidence; Scopelliti et al., 2015; Ybarra, 2018). Recent research suggests that more skilled decision makers (i.e., more numerate) may often use their skills to develop a more integrated representative understanding of the relevant information, which typically allows them to better monitor their own understanding (e.g., more calibrated self-evaluations and less overconfidence bias; Cokely et al., 2018; Ghazal et al., 2014; Ybarra, 2018). To test these potential relations, Study 2 included a measure of confidence in one's casual comprehension, which was used in concert with the causal misunderstanding item to include overconfidence in the cognitive model. Following previous research, I hypothesized that overconfidence would at least partially mediate the relations between numeracy, scientific reasoning, risk perceptions, and judgment biases.

Participants and Procedure

The data were collected via Qualtrics in Fall 2023 using an undergraduate student sample. The study was designed to take roughly 30 minutes to complete, and the average completion time was 20 minutes. Out of 400 total cases, 342 were used for the analyses after excluding participants who took insufficient time (i.e., those who took less than 10 minutes to complete the study). Of the 342 participants, 269 (78.6%) identified as female, 71 (20.8%) identified as male, and 2 (0.6%) identified as non-binary. Participants reported ages ranged between 18 and 34 years old.

Participants read two news article excerpts in a randomized order (both excerpts were also used in Study 1; Cosdon, 2022; Kissell, 2022). Each excerpt reported scientific findings published in peer-reviewed journals and presented an opportunity to make a causal theory error. After each excerpt, participants answered a question related to causal misunderstanding (i.e., whether they thought the correlational information presented in the article was causal), confidence, risk perceptions (i.e., how concerned they think someone should be about the risk), and multiple judgments about behavioral intentions following the protocol in Study 1.

Measures

All measures from Study 1 were included in Study 2. This includes: (i) *Statistical Numeracy*, (ii) *Scientific Reasoning*, (iii) *Causal Misunderstanding*, (iv) *Risk Perceptions*, and (v) *Judgment Biases*. However, in this study, causal misunderstanding was measured by a single binary item (True/False) and one additional variable was measured: *Confidence*.

Causal Misunderstanding

Causal misunderstanding was measured by one item which asked participants whether they thought the information presented in the article was causal (e.g., *According to the article, watching a lot of TV causes dementia*). Respondents were required to evaluate the statement as True or False. Correct answers (False) were scored as “1”, while incorrect answers (True) were scored as “0”.

Confidence

Confidence was measured with an item asking the degree to which participants believe their causal judgment was accurate (i.e., *How confident are you that your above judgment is accurate?*). The scale ranged from 0% confident to 100% confident in increments of 10%.

Causal Misunderstanding: Overconfidence

The score for this variable was calculated as a participant's causal misunderstanding score subtracted from their confidence score. Confidence scores were converted from whole numbers to decimals for the calculation (e.g., 100% confident = 1.0, 90% confident = 0.9). Scores for this variable ranged from -1 (highly underconfident) to 1 (highly overconfident), where a score of 0 indicated neither over nor underconfident. For example, someone who scores a 0 on the causal misunderstanding item (incorrect) but is 90% confident that their answer was correct would receive a score of 0.9 (i.e., 0.9 – 0 = 0.9), indicating a high degree of overconfidence in their judgment accuracy.

Results and Analyses

All statistical analyses were completed in R version 4.3.0. Base R packages were used to compute descriptive statistics and correlations for all Study 2 variables (Table 3). To test relationships between numeracy, scientific reasoning, causal misunderstanding,[6] and downstream judgment biases, a structural modeling approach was again used, and a latent trait of judgment biases was estimated in the structural model (Figure 2). Relevant indirect effects within the model were tested and presented with 95% CI estimated using 1,000 bootstrap samples (Table 4).

[6] *Causal misunderstanding*, as discussed in the results, refers to the combined measure of participants' causal misunderstanding score and confidence score (e.g., higher scores demonstrate overconfidence as a function of accuracy of understanding; see Method section "Causal Misunderstanding: Overconfidence" for details on variable calculation).

Table 3

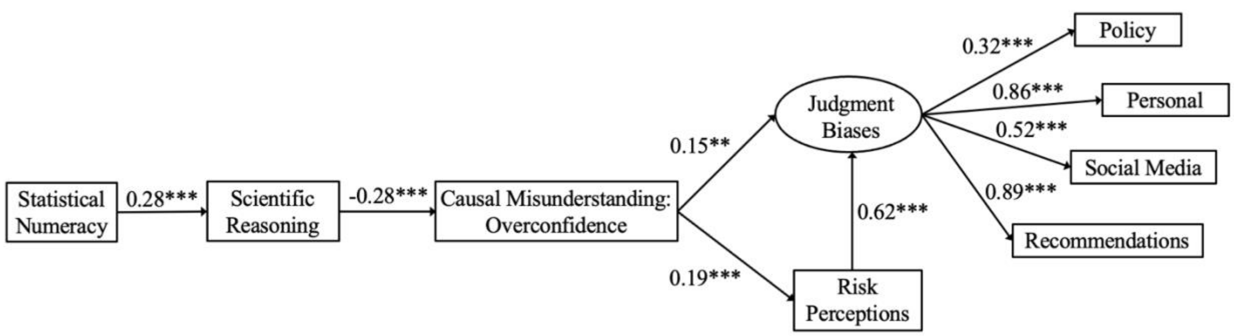
Descriptive Statistics and Correlations for Study 2 Variables

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. BNT-S	3.72	1.74								
2. SRS	5.85	2.51	.28**							
3. Causal Misunderstanding	0.29	0.43	-.12*	-.28**						
4. Risk Perceptions	3.88	1.03	.03	-.11*	.19**					
5. Judgment Biases	3.63	0.79	.05	-.23**	.26**	.55**				
6. Social Media	2.83	1.19	-.02	-.27**	.21**	.31**	.73**			
7. Recommendations	4.21	1.02	.07	-.14*	.23**	.60**	.83**	.43**		
8. Personal	4.10	1.05	.12*	-.12*	.23**	.53**	.85**	.49**	.77**	
9. Policy	3.39	0.97	-.04	-.13*	.09	.22**	.56**	.17**	.27**	.28**

Note. *$p < .05$, **$p < .01$.

Figure 2

Structural Equation Model with Overconfidence Calculated as a Function of Judgment Accuracy



Note. Standardized coefficients are reported. *$p < .05$, **$p < .01$, ***$p < .001$. Model estimated using the *lavaan* package in R (Rosseel, 2012).

As shown in Figure 2, structural equation modeling indicated numeracy was related to reduced causal misunderstanding because numerate people tended to be better at scientific reasoning (i.e., less overconfidence as a function of accuracy of understanding). The model exhibited acceptable fit, $\chi^2_{19} = 45.40$, $p = 0.001$, CFI = 0.96, TLI = 0.94, RMSEA = 0.06 [0.04, 0.09], SRMR = 0.06, such that statistical numeracy predicted scientific reasoning ($\beta = 0.28$), which in turn predicted causal misunderstanding ($\beta = -0.28$). Additionally, structural equation modeling revealed that the improved causal misunderstanding variable fully mediated the relationship between scientific reasoning and all downstream judgment biases (see Figure 2). As shown in Table 4, an analysis of indirect effects confirmed that numeracy exerted a significant indirect effect on causal misunderstanding (-.08, 95% CI [-.12, -.04]), and risk perceptions (-.02, 95% CI [-.03, -.01]). Further, numeracy exerted a significant indirect effect on judgment biases, through scientific reasoning, causal misunderstanding, and risk perceptions (-.02, 95% CI [-.04, -.01]).

Table 4

Total Indirect Effects (Standardized Coefficients) for Key Variables in Study 2

Paths	Estimate	SE	Bootstrapped 95% CI
(1) Numeracy → scientific reasoning → **causal misunderstanding**	-.08***	.02	[-.12, -.04]
(1) Numeracy → scientific reasoning → causal misunderstanding → **risk perceptions**	-.02*	.01	[-.03, -.01]
(1) Numeracy → scientific reasoning → causal misunderstanding → **judgment biases** (2) Numeracy → scientific reasoning → causal misunderstanding → risk perceptions → **judgment biases**	-.02**	.01	[-.04, -.01]

Note. Estimate reflects the total indirect effect from multiple paths for each dependent variable. Effects were estimated with 1,000 bootstrap samples. SE = standard error; CI = confidence interval. *p < .05, **p < .01, ***p < .001.

Study 2 Discussion

Study 2 verified and extended findings from Study 1 using refined assessments, including a metacognitive judgment on the causal theory error question. Consistent with primary findings from Study 1, results from Study 2 again suggested that more numerate people were less likely to misinterpret correlational information as causal, and in turn, were less likely to exhibit downstream judgment biases. Once again, the structural equation model indicated that the benefits of numeracy primarily resulted because numerate people had previously acquired specialized scientific reasoning skills, which fully explained the relations between numeracy and downstream judgment biases. By introducing a new variable (i.e., causal misunderstanding: overconfidence), Study 2 also provided novel evidence on the role of a metacognitive factor—i.e., overconfidence in one's own understanding of regular science reporting. As anticipated, overconfidence in science report comprehension fully mediated the relationship between scientific reasoning and all downstream judgment biases (i.e., people who were more overconfident in their understanding were more likely to express judgment biases).

Chapter 4: General Discussion

The current findings suggest numerate people may typically be more likely to accurately evaluate and better understand at least some important elements of regular science reporting (i.e., less likely to mistake correlational evidence as causal). These studies also appear to be the first to test a cognitive model of people's interpretations of regular science reporting, linking numeracy, scientific reasoning, causal theory errors, and downstream judgment biases. Findings indicate that numerate people may be less likely to make causal theory errors when reviewing regular science reporting primarily because numerate people are more likely to have acquired and applied specialized scientific reasoning skills. Study 2 replicated and extended these findings, further revealing that numerate people's scientific reasoning skills also helped them gain more metacognitive insight into their own comprehension (i.e., reduced overconfidence). In turn, numerate people's ability to accurately evaluate and understand the regular science reporting promoted more informed and less biased downstream judgments (e.g., risk perceptions, behavioral intentions, policy preferences). Overall, results reveal some potential obstacles that may make independent (unaided) evaluation of regular science reporting difficult for many people. Results also shed new light on one way numerate individuals may independently acquire some of the knowledge they typical rely on to make more informed and less biased judgments and decisions.

Risk Literacy is Not (Just) Statistical Numeracy

The theoretical construct of *Risk Literacy* has been defined as "the ability to evaluate and understand risk" (Cokely et al., 2018; see also Cokely et al., 2012; 2013; 2014; Feltz et al., 2022; Feltz & Cokely, *in press*; Garcia-Retamero & Cokely, 2017; Gigerenzer, 2012; 2015; Reyna & Brainerd, 2023). This conceptualization suggests that risk literacy skills are not identical to statistical numeracy skills, even though the two sets of skills tend to be robustly linked (e.g., statistical numeracy skills fundamentally involve practical skills for probabilistic and inductive reasoning; Cokely et al., 2012; 2013; 2014; 2018). A related conceptual view has recently been advanced by Aven (2023) suggesting that risk literacy should not be too narrowly construed as probabilistic reasoning.[7] Instead, Aven argues for a broader concept of risk literacy that includes other skills present in practitioners, scientists, and experts in fields of risk analysis, risk science, and decision analysis. For example, Aven noted that a broader view of risk literacy should include skills related to other fundamental concepts of science (e.g., distinguishing correlation from causation; Aven, 2020), which are often elements of expert risk analysis. Accordingly, the current studies provide some of the first evidence linking both numeracy and science reasoning with risk literacy and skilled decision making more generally.

The current findings also accord with the *Risk Literacy Skills for Knowledge Hypothesis* that was proposed in this thesis, which holds that numerate people may generally be more likely to independently acquire a broad constellation of skills, including specialized scientific reasoning skills that help give rise to risk literacy and risk knowledge acquisition. This hypothesis is also consistent with evidence indicating that numerate people tend to independently acquire more advanced graph literacy skills, which also promote risk literacy in ways that mediate the influence of numeracy (Garcia-Retamero & Cokely, 2013; 2017). However, unlike most relations between statistical numeracy and graph literacy, research suggests numeracy skills may be prerequisites for the acquisition of science reasoning skills (e.g., necessary but not sufficient for effective scientific reasoning; National Academies of Sciences, 2016). If correct, this may have notable implications

[7] Aven suggested the concept of risk literacy should not be conflated with probabilistic reasoning, rejecting the notion that risk literacy should be primarily defined by the Knightian dichotomy between risk and uncertainty (e.g., where risk is strictly viewed as known, objective probabilities; Knight, 1921).

for the order of training programs used to develop risk literacy skills (e.g., the development of a broad foundation in statistical numeracy should likely come before science reasoning training).

Limitations

One limitation of the current investigation is that it employed a relatively narrow band of materials and participants. As such, future investigations may want to consider testing a wider range of articles (e.g., other publishers and formats, more and less controversial or familiar), across multiple domains (e.g., other than health related domains; Dhami et al., 2004). Researchers may also want to consider recruiting more representative and diverse samples (e.g., other than US public college students). That said, given the nature of the mechanisms involved in the current research, it seems likely findings will generalize to a meaningful extent (i.e., key findings are primarily about the configural relations between skills, reasoning, and interpretations). Still, it is possible that other factors such as prior knowledge[8] or biases could interact in unusual ways. For example, some evidence suggests that motivated reasoning biases might sometimes overshadow the benefits of numeracy or science reasoning in special situations, such as those dealing with controversial risks (e.g., guns, climate change) that involve entrenched biases, extreme worldviews, and other conflicts of interest (Cho et al., 2023; Kahne & Bowyer, 2017; Roh et al., 2015; Taber & Lodge, 2006).

Another potential limitation of the current investigation is that the direct effect of numeracy on causal misunderstanding was not significant in Study 1. The lack of a significant direct relation in Study 1 appears to be explained by limits of the psychometric sensitivity of the

[8] Prior knowledge was not measured in the current studies and theoretically there may be other campaigns and research findings that may be relevant (e.g., Ihira et al., 2022; Inoue et al., 2023; Vallance et al., 2018). However, to limit the role of prior knowledge, the causal judgment and behavioral intentions questions explicitly asked participants to respond solely based on the information provided in the brief article excerpts. However, it may still be the case that participants utilized prior knowledge, which may have shaped their responses to causal judgments and downstream behavioral intentions.

misunderstanding variable that was used. The sensitivity issue was addressed in Study 2 by modifying the misunderstanding variable, which then demonstrated a small but significant direct effect between numeracy and causal theory errors. The indirect effect linking numeracy to causal theory errors via scientific reasoning skills was also significant in both Study 1 and 2. Taken together, the overall direct effect of numeracy appears to be modest. Still, even if the relationship only confers very small benefits on each specific science report, these benefits seem likely to add up over time. Given the evidence that overall engagement with science news has increased in recent years (Hendriks et al., 2020; National Science Board, 2018; Pavelle & Wilkinson, 2020; Saks & Tyson, 2022), if numerate people acquire just one extra fact from each of the articles they read, numerate readers would be likely to learn thousands of additional decision-relevant facts each year.

Risk Literacy Difficulty Analysis

Given the modest yet robust link between numeracy, scientific reasoning, and causal misunderstanding that was observed, I conducted a Risk Literacy Difficulty Analysis to further investigate potential practical implications of these relations. Using the methodology developed by Allan (2021)[9], I estimated (a) how difficult each task was in terms of the overall Risk Literacy Difficulty Level (i.e., an estimated standardized theta score in a unidimensional latent trait space), (b) the proportion of US adults likely to misunderstand similar science reports, and (c) how much numeracy skill would typically be needed to likely avoid misinterpretations of science reports like those used in the current studies. In other words, this methodology provides a means to estimate and compare the Risk Literacy Difficulty Level of different tasks, which can then be transformed

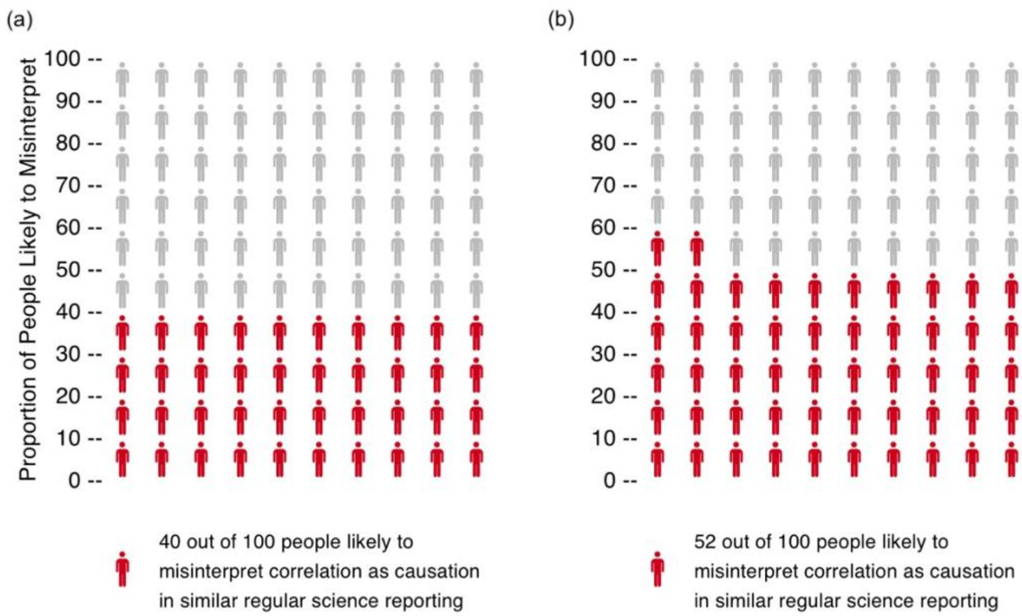
[9] Although estimating accuracy is known to depend on multiple model assumptions and research design factors (e.g., how representative are tasks; how psychometrically reliable are dependent measures), initial investigations of the Risk Literacy Difficulty Analysis method indicate it may often have robust predictive validity, as evidenced in out-of-sample cross-validation tests of the method.

to estimate the Probability of Misunderstanding among the general US adult population, as well as a means to estimate the Minimum Numeracy Score threshold (i.e., the minimum score on the Berlin Numeracy Test typically required for someone to be "more likely than not" to avoid errors and biases, defined as achieving at least 50% task accuracy; Allan, 2021).

As such, using data from Study 2, I conducted a Risk Literacy Difficulty Analysis for each of the two article excerpts (i.e., TV and dementia; Vitamin D and COVID-19 severity) using participants' BNT-S and causal misunderstanding scores. Results from the analysis revealed that the tasks were relatively difficult overall. When evaluating the TV and dementia article, the estimated minimum BNT-S score for accurate causal interpretation (i.e., to correctly reject the causal claim) was 4.4 out of 7. Additionally, the analysis indicates that around 40 percent of the general US adult population may be likely to make casual theory errors when evaluating similar science reports (see Figure 3). Likewise, when applied to the Vitamin D and COVID-19 severity article, results indicated that the minimum score on the BNT-S at which a person is likely to make an accurate causal interpretation was about 4.8 (out of 7), suggesting that around 52 percent of the general US adult population may be likely to misunderstand correlational information as causal when evaluating similar regular science reports. Overall, the Risk Literacy Difficulty Analysis suggests that nearly half of all adult residents in the United States may be likely to misunderstand essential information in regular science reports like the ones used in the current study. If correct, it seems likely that many people are likely to misinterpret and be misinformed by regular science reports.

Figure 3

Risk Literacy Difficulty Levels for (a) the TV and Dementia article and (b) the Vitamin D and COVID-19 Severity article



Conclusions

The current work sheds light on how numerate people become more informed and less biased decision makers. Consistent with the *Risk Literacy Skills for Knowledge Hypothesis*, the findings suggest that it may be because they (i) tend to independently acquire specialized scientific reasoning skills that promote the ability to evaluate and understand risk information (i.e., risk literacy), and (ii) use these relevant risk literacy skills to both learn from and avoid being misled by regular science reporting. However, results from the current studies suggest there may be a disparity in the ability to understand science information: for roughly half of all US adults, information provided in regular science reporting may be more confusing than it is *informative*. One way to address this disparity may be to improve science reporting standards (e.g., how researchers and journalists report findings). However, in light of previous and current research that suggests deficits in relevant skills and knowledge may lead to misunderstanding even transparent scientific reporting, future research should also aim to develop specialized risk literacy skills training programs that empower diverse individuals to accurately interpret and utilize new science information and complex risk communications across domains. Theoretically, these training programs would equip people with skills that reduce misunderstanding for regular science reporting (e.g., numeracy, science reasoning), and may also protect against biased reasoning even when faced with information that is not transparent (e.g., overexaggerated claims in media reports). Taken together, results from the current studies suggest that Risk Literacy Knowledge is Power. When equipped with the skills to accurately evaluate and understand risk information, people are more likely to become more informed and less biased decision makers.

References

Adams, R. C., Sumner, P., Vivian-Griffiths, S., Barrington, A., Williams, A., Boivin, J., ... & Bott, L. (2017). How readers understand causal and correlational expressions used in news headlines. *Journal of experimental psychology: applied*, *23*(1), 1.

Aven, T. (2020). Risk science contributions: three illustrating examples. *Risk Analysis*, *40*(10), 1889-1899.

Aven, T. (2023). Risk literacy: Foundational issues and its connection to risk science. *Risk Analysis*.

Allan, J. N., Ripberger, J. T., Ybarra, V. T., & Cokely, E. T. (2017, June). Tornado risk literacy: Beliefs, biases, and vulnerability. In *13th Bi-Annual Int. Conf. on Naturalistic Decision Making and Uncertainty* (pp. 284-290). Bath, United Kingdom: University of Bath.

Allan, J. N. (2021). Statistical numeracy norms and decision vulnerability benchmarks: A norm-referenced method for estimating the risk literacy difficulty level of choices and risk communications. [Doctoral Dissertation, University of Oklahoma].

Bao, L., Koenig, K., Xiao, Y., Fritchman, J., Zhou, S., & Chen, C. (2022). Theoretical model and quantitative assessment of scientific thinking and reasoning. *Physical Review Physics Education Research*, *18*(1), 010115.

Barz, D. L., & Achimaş-Cadariu, A. (2016). The development of scientific reasoning in medical education: a psychological perspective. *Clujul medical*, *89*(1), 32.

Bašnáková, J., Čavojová, V., & Šrol, J. (2021). Does concrete content help people to reason scientifically? Adaptation of Scientific reasoning scale. *Science & Education*, *30*(4), 809-826.

Bott, L., Bratton, L., Diaconu, B., Adams, R. C., Challenger, A., Boivin, J., ... & Sumner, P. (2019). Caveats in science-based news stories communicate caution without lowering interest. *Journal of Experimental Psychology: Applied*, *25*(4), 517.

Bleske-Rechek, A., Morrison, K. M., & Heidtke, L. D. (2015). Causal inference from descriptions of experimental and non-experimental research: Public understanding of correlation-versus-causation. *The Journal of general psychology*, *142*(1), 48-70.

Bratton, L., Adams, R. C., Challenger, A., Boivin, J., Bott, L., Chambers, C. D., & Sumner, P. (2019). The association between exaggeration in health-related science news and academic press releases: a replication study. *Wellcome open research*, *4*.

Bratton, L., Adams, R. C., Challenger, A., Boivin, J., Bott, L., Chambers, C. D., & Sumner, P. (2020). Causal overstatements reduced in press releases following academic study of health news. *Wellcome Open Research*, *5*.

Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*(2), 59-69.

Bruine de Bruin, W., & Bostrom, A. (2013). Assessing what to address in science communication. *Proceedings of the National Academy of Sciences*, *110*(supplement_3), 14062-14068.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology*, *4*(1), 55-81.

Cho, J., Cokely, E. T., Ramasubramanian, M., Allan, J. N., Feltz, A., & Garcia-Retamero, R. (2023). Numeracy does not polarize climate change judgments: Numerate people are more knowledgeable and knowledge is power. *Decision*.

Cokely, E. T., Cho, J., Allan, J. N., Ramasubramanian, M., Feltz, A., & Garcia-Retamero, R. (2025). Risk literacy knowledge is power: On the dominant role of representative understanding in skilled decision making [Manuscript in preparation]. In M. Williams, A. Kozbelt, F. Preckel, & R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance.* Cambridge University Press.

Cokely, E., Feltz, A., Ghazal, S., Allan, J., Petrova, D., & Garcia-Retamero, R. (2018). Skilled Decision Theory: From Intelligence to Numeracy and Expertise. In K. Ericsson, R. Hoffman, A. Kozbelt, & A. Williams (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge Handbooks in Psychology, pp. 476-505). Cambridge: Cambridge University Press.

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

Cokely, E. T., Ghazal, S., & Garcia-Retamero, R. (2014). Measuring numeracy. *Numerical reasoning in judgments and decision making about health*, 11-38.

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.

Cosdon, N. (2022, February 3). *Vitamin D Deficiency Increases Risk of Severe or Fatal COVID-19.* ContagionLive. https://www.contagionlive.com/view/vitamin-d-deficiency-increases-risk-of-severe-or-fatal-covid-19

Cram, P., Fendrick, A. M., Inadomi, J., Cowen, M. E., Carpenter, D., & Vijan, S. (2003). The impact of a celebrity promotional campaign on the use of colon cancer screening: the Katie Couric effect. *Archives of Internal Medicine*, *163*(13), 1601-1605.

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological bulletin*, *130*(6), 959.

Díaz, C., Dorner, B., Hussmann, H., & Strijbos, J. W. (2023). Conceptual review on scientific reasoning and scientific thinking. *Current Psychology*, *42*(6), 4313-4325.

Drummond, C., & Fischhoff, B. (2017). Development and validation of the scientific reasoning scale. *Journal of Behavioral Decision Making*, *30*(1), 26-38.

Drummond, C., & Fischhoff, B. (2019). Does "putting on your thinking cap" reduce myside bias in evaluation of scientific evidence?. *Thinking & Reasoning*, *25*(4), 477-505.

Dunbar K., & Fugelsang, J. Scientific Thinking and Reasoning. In: Holyoak KJ, Morrison RG, editors. *The Cambridge Handbook of Thinking and Reasoning.* Vol. 705. Cambridge University Press; 2005. p. 720.

Elliott, B. (2022, March 2). *Just 30-90 minutes of resistance training weekly decreases risk of premature death – new research.* The Conversation. https://theconversation.com/just-30-90-minutes-of-resistance-training-weekly-decreases-risk-of-premature-death-new-research-178209

Ericsson, K. A., Charness, N., Hoffman, R. R., & Feltovich, P. J. (eds.) (2006). The Cambridge handbook of expertise and expert performance. Cambridge University Press.

Ericsson, K. A., Hoffman, R. R., Kozbelt, A., & Williams, A. M. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard business review*, *85*(7/8), 114.

Feltz, A. (2016). Financial surrogate decision making: Lessons from applied experimental philosophy. *The Spanish Journal of Psychology*, *19*, E56.

Feltz, A., Caton, J., Cogley, Z., Engel Jr, M., Feltz, S., Ilea, R., ... & Offer-Westort, T. (2023). Using food frequency questionnaires to measure traits: A case study of human consumption of animals and animal products. *Psychology of Human-Animal Intergroup Relations*, *2*, 1-22.

Feltz, A., & Cokely, E. T. (in press). *Diversity and Disagreement: From Fundamental Biases to Ethical Interactions.* Palgrave Macmillan.

Feltz, A., Tanner, B., Hoang, G., Holt, J., & Muhammad, A. (2022). Free Will and Skilled Decision Theory. *Advances in Experimental Philosophy of Free Will and Responsibility*, 185.

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., ... & Eberle, J. (2014). Scientific reasoning and argumentation: advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, *2*(3), 28-45.

Funk, C., Gottfried, J., & Mitchell, A. (2017, September 20). *A majority of Americans rely on general outlets for science news but more say specialty sources get the facts right about science.* Pew Research Center. https://www.pewresearch.org/journalism/2017/09/20/science-news-and-information-today/

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.

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). Using visual aids to help people with low numeracy make better decisions. *Numerical reasoning in judgments and decision making about health*, 153-174.

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.

Garcia-Retamero, R., & Cokely, E. T. (2017). Designing visual aids that promote risk literacy: A systematic review of health research and evidence-based design heuristics. *Human factors*, *59*(4), 582-627.

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

Garcia-Retamero, R., & Hoffrage, U. (2006). How causal knowledge simplifies decision-making. *Minds and Machines*, *16*, 365-380.

Garcia-Retamero, R., Petrova, D., Cokely, E. T., & Joeris, A. (2020). Scientific risk reporting in medical journals can bias expert judgment: Comparing surgeons' risk comprehension across reporting formats. *Journal of Experimental Psychology: Applied*, *26*(2), 283.

Garcia-Retamero, R., Sobkow, A., Petrova, D., Garrido, D., & Traczyk, J. (2019). Numeracy and risk literacy: What have we learned so far?. *The Spanish journal of psychology*, *22*, E10.

Garcia-Retamero, R., Wallin, A., & Dieckmann, A. (2007). Does causal knowledge help us be faster and more frugal in our decisions?. *Memory & cognition*, *35*, 1399-1409.

Garrido, D., Petrova, D., Cokely, E., Carballo, G., & Garcia-Retamero, R. (2021). Parental risk literacy is related to quality of life in spanish families of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *51*(7), 2475-2484.

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. (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. (2015). Risk savvy: How to make good decisions. New York: Penguin.

Grounds, M. A., & Joslyn, S. L. (2018). Communicating weather forecast uncertainty: Do individual differences matter?. *Journal of experimental psychology: applied*, *24*(1), 18.

Hendriks, F., Mayweg-Paus, E., Felton, M., Iordanou, K., & Jucks, R. (2020). Constraints and affordances of online engagement with scientific information—A literature review. *Frontiers in Psychology*, *11*, 572744.

Hilton, A., & Hilton, G. (2016). Proportional reasoning: An essential component of scientific understanding. *Teaching Science*, *62*(4), 32-42.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Ihira, H., Sawada, N., Inoue, M., Yasuda, N., Yamagishi, K., Charvat, H., ... & Tsugane, S. (2022). Association between physical activity and risk of disabling dementia in Japan. *JAMA Network Open*, *5*(3), e224590-e224590.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures* (Vol. 22). Psychology Press.

Inoue, K., Tsugawa, Y., Mayeda, E. R., & Ritz, B. (2023). Association of daily step patterns with mortality in US adults. *JAMA Network Open*, *6*(3), e235174-e235174.

Kahne, J., & Bowyer, B. (2017). Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation. *American educational research journal*, *54*(1), 3-34.

Kambeyo, L., & Csapó, B. (2018). Scientific Reasoning skills: A theoretical Backgrounds to science education. In *Reform Forum: Journal for Educational Research in Namibia* (Vol. 26, No. 1, pp. 27-36).

Kissell, C. (2022, August 26). *Watching too much TV linked to increased risk of dementia*. Money Talks News. https://www.moneytalksnews.com/watching-too-much-tv-linked-to-increased-risk-of-dementia/

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive science*, *12*(1), 1-48.

Klein, G. A. (1999). Sources of power: How people make decisions. MIT Press.

Knight, F. H. (1921). *Risk, uncertainty and profit* (Vol. 31). Houghton Mifflin.

Kuhn, D., & Dean Jr, D. (2004). Connecting scientific reasoning and causal inference. *Journal of cognition and development*, *5*(2), 261-288.

Kuhn, D. (2012). The development of causal reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 327-335.

Matthews, A., Herrett, E., Gasparrini, A., Van Staa, T., Goldacre, B., Smeeth, L., & Bhaskaran, K. (2016). Impact of statin related media coverage on use of statins: interrupted time series analysis with UK primary care data. *BMJ*, *353*.

Michal, A. L., & Shah, P. (2024). A Practical Significance Bias in Laypeople's Evaluation of Scientific Findings. *Psychological Science*.

Miller, J. D. (1998). The measurement of civic scientific literacy. *Public understanding of science, 7*(3), 203.

Ministry of Education. (1959). *15 to 18: A Report of the Central Advisory Council for Education (England) Volume 1 and 2*. HM Stationery Office.

National Academies of Sciences, Engineering, and Medicine. (2016). Science literacy: Concepts, contexts, and consequences.

National Academies of Sciences, Division of Behavioral, Committee on the Science of Science Communication, & A Research Agenda. (2017). Communicating science effectively: A research agenda.

National Science Board. (2018). *Science and engineering indicators 2018*. National Science Foundation.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-hall.

Norris, S. P., Phillips, L. M., & Korpan, C. A. (2003). University students' interpretation of media reports of science and its relationship to background knowledge, interest, and reading difficulty. *Public Understanding of Science, 12*(2), 123-145.

Okan, Y., Galesic, M., & Garcia-Retamero, R. (2016). How people with low and high graph literacy process health graphs: Evidence from eye-tracking. *Journal of Behavioral Decision Making*, *29*(2-3), 271-294.

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.

Opitz, A., Heene, M., & Fischer, F. (2017). Measuring scientific reasoning–a review of test instruments. *Educational Research and Evaluation*, *23*(3-4), 78-101.

Pavelle, S., & Wilkinson, C. (2020). Into the digital wild: Utilizing Twitter, Instagram, YouTube, and Facebook for effective science and environmental communication. *Frontiers in Communication*, *5*, 575122.

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

Peters, E. (2020). *Innumeracy in the wild: Misunderstanding and misusing numbers*. Oxford University Press.

Peters, E., & Bunquin, J. B. (2024). The Power of Numeric Evidence in Science Communication. *Current Directions in Psychological Science*, 09637214241242465.

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., Cokely, E. T., Sobkow, A., Traczyk, J., Garrido, D., & Garcia-Retamero, R. (2023). Measuring feelings about choices and risks: The Berlin emotional responses to risk instrument (BERRI). *Risk Analysis*, *43*(4), 724-746.

Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E., Heredia Carrasco, A., Arrebola Moreno, A., & Ramírez Hernández, J. A. (2017). Numeracy predicts risk of pre-hospital decision delay: A retrospective study of acute coronary syndrome survival. *Annals of Behavioral Medicine*, *51*(2), 292-306.

Petrova, D., Garcia-Retamero, R., Catena, A., & Van der Pligt, J. (2016). To screen or not to screen: What factors influence complex screening decisions?. *Journal of Experimental Psychology: Applied*, *22*(2), 247.

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., Kostopoulou, O., Delaney, B. C., Cokely, E. T., & Garcia-Retamero, R. (2018). Strengths and gaps in physicians' risk communication: a scenario study of the influence of numeracy on cancer screening communication. *Medical Decision Making*, *38*(3), 355-365.

Petrova, D., Mas, G., Navarrete, G., Rodriguez, T. T., Ortiz, P. J., & Garcia-Retamero, R. (2019). Cancer screening risk literacy of physicians in training: an experimental study. *PLoS One*, *14*(7), e0218821.

Ramasubramanian, M., Allan, J. N., Retamero, R. G., Jenkins-Smith, H., & Cokely, E. T. (2019, November). Flood risk literacy: Communication and implications for protective action. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 63, No. 1, pp. 1629-1633). Sage CA: Los Angeles, CA: SAGE Publications.

Reyna, V. F., & Brainerd, C. J. (2023). Numeracy, gist, literal thinking and the value of nothing in decision making. *Nature Reviews Psychology*, *2*(7), 421-439.

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.

Roh, S., McComas, K. A., Rickard, L. N., & Decker, D. J. (2015). How motivated reasoning and temporal frames may polarize opinions about wildlife disease risk. *Science Communication*, *37*(3), 340-370.

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., ... & Van Der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science*, *7*(10), 201199.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of statistical software*, *48*, 1-36.

Saks, E., & Tyson, A. (2022, November 10). *Americans report more engagement with science news than in 2017*. Pew Research Center. https://www.pewresearch.org/short-reads/2022/11/10/americans-report-more-engagement-with-science-news-than-in-2017/

Savadori, L., Dickson, M. M., Micciolo, R., & Espa, G. (2022). The polarizing impact of numeracy, economic literacy, and science literacy on the perception of immigration. *Plos one*, *17*(10), e0274680.

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.

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., & Kassam, K. S. (2015). Bias blind spot: Structure, measurement, and consequences. *Management Science*, *61*(10), 2468-2486.

Seifert, C. M., Harrington, M., Michal, A. L., & Shah, P. (2022). Causal theory error in college students' understanding of science studies. *Cognitive Research: Principles and Implications*, *7*(1), 4.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of educational psychology*, *89*(2), 342.

Stryker, J. E., Moriarty, C. M., & Jensen, J. D. (2008). Effects of newspaper coverage on public knowledge about modifiable cancer risks. *Health communication*, *23*(4), 380-390.

Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C. A., Davies, A., ... & Chambers, C. D. (2014). The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ*, *349*.

Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Bott, L., Adams, R., ... & Chambers, C. D. (2016). Exaggerations and caveats in press releases and health-related science news. *PloS one*, *11*(12), e0168217.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American journal of political science*, *50*(3), 755-769.

Tanner, B., & Feltz, A. (2022). Comparing effects of default nudges and informing on recycled water decisions. *Journal of Experimental Psychology: Applied*, *28*(2), 399.

Tanner, B., McDonald, P., Mahmoud-Elhaj, D., Sabatini, D., Chamberlain, J. F., Nijhawan, A., ... & Feltz, A. (2024). Not All Information is Informative: An Exploration of Educational Content on Recycled Potable Water Knowledge and Acceptance. *Water Research*, 121254.

Vallance, J. K., Gardiner, P. A., Lynch, B. M., D'Silva, A., Boyle, T., Taylor, L. M., ... & Owen, N. (2018). Evaluating the evidence on sitting, smoking, and health: is sitting really the new smoking?. *American journal of public health, 108*(11), 1478-1482.

Van der Linden, S. (2023). *Foolproof: Why misinformation infects our minds and how to build immunity.* WW Norton & Company.

Voss, N. M. (2024). The effects of careless responding on the fit of confirmatory factor analysis and item response theory models. *Behavior Research Methods*, *56*(2), 577-599.

Woloshin, S., Schwartz, L. M., Casella, S. L., Kennedy, A. T., & Larson, R. J. (2009). Press releases by academic medical centers: not so academic?. *Annals of Internal Medicine*, *150*(9), 613-618.

Woodyatt, A. (2021, September 22). *People who eat more dairy fat have lower risk of heart disease, study suggests*. CNN. https://www.cnn.com/2021/09/22/health/dairy-cardiovascular-disease-intl-scli-wellness-scn/index.html#:~:text=People%20who%20eat%20more%20dairy,heart%20disease%2C%20study%20suggests%20%7C%20CNN

Xiong, C., Shapiro, J., Hullman, J., & Franconeri, S. (2019). Illusion of causality in visualized data. *IEEE transactions on visualization and computer graphics*, *26*(1), 853-862.

Ybarra, V., Allan, J., Garcia-Retamero, R., Feltz, A., & Cokely, E. T. (2018). Rethinking the bias blind spot: Numerate people are less biased and they know it. *Society for Judgment and Decision Making*.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental review*, *20*(1), 99-149.

Appendix

Article 1 (Included in Studies 1 & 2)[10]

Watching Too Much TV Linked to Increased Risk of Dementia

With age comes the temptation to engage in watching hours of TV and other passive activities. But a new study suggests that such passivity might put our brains at risk.

Being 60 years of age or older and engaging in long periods of sitting passively may increase the risk of developing dementia, according to researchers at the University of Southern California and the University of Arizona.

However, the risk diminishes for those who sit while actively engaging the mind, such as when using the computer. The findings were published in the journal Proceedings of the National Academy of Science.

In a press release summarizing the study findings, study author David Raichlen, professor of biological sciences and anthropology at USC, says: “We know from past studies that watching TV involves low levels of muscle activity and energy use compared with using a computer or reading. And while research has shown that uninterrupted sitting for long periods is linked with reduced blood flow in the brain, the relatively greater intellectual stimulation that occurs during computer use may counteract the negative effects of sitting.”

The researchers analyzed self-reported data from the U.K. Biobank, a large biomedical database of participants throughout the United Kingdom. Specifically, they looked at data on more than 145,000 participants who were age 60 or older and were followed for about 12 years, on average.

The researchers found that spending time watching TV was associated with an increased risk of dementia even among people who are physically active. By contrast, spending time using a computer was associated with a reduced risk of developing dementia.

[10] Study 1 included a between subjects experimental manipulation with two conditions. Participants in the “risk” condition were presented these 4 articles exactly the way they appear in the Appendix, while those in the “correlation” condition were shown these articles with slightly modified wording. Results were collapsed as there was no significant effect of condition.

Article 2 (Included in Studies 1 & 2)

Vitamin D Deficiency Increases Risk of Severe or Fatal COVID-19

Vitamin D is crucial to maintain bone health and facilitate immune system function. Early in the pandemic, health officials encouraged the public to take vitamin D supplements to boost immune response and potentially protect against COVID-19.

Now, studies are increasingly demonstrating that low vitamin D levels pose an increased risk of COVID-19 infection and death. One study, published today in PLOS ONE, is among the first to analyze vitamin D levels prior to COVID-19 infection.

Jointly led by Bar-Ilan University and Galilee Medical Center, the study looked for any correlation between pre-infection serum 25-hydroxyvitamin D (25(OH)D) level and COVID-19 severity and mortality. Investigators analyzed hospital records of individuals admitted to the Galilee Medical Center in Nahariya, Israel from April 7, 2020-February 4, 2021. Included subjects had a positive PCR COVID-19 test and 25(OH)D levels documented 14-730 days prior to their infection.

The investigators grouped patients by disease severity and level of 25(OH)D. Vitamin D levels were categorized as deficient, insufficient, adequate, or high-normal. They determined COVID-19 disease severity using multivariable regression analysis. Utilizing a cosinor model, the investigators isolated the influence of a sinusoidal pattern of seasonal 25(OH)D fluctuations.

Of the 1176 patients included in the study, 253 had a record of their 25(OH)D level before COVID-19 infection. Patients with a vitamin D deficiency, defined as < 20 ng/mL, were 14 times more likely to have severe or fatal COVID-19 disease than patients with 25(OH)D ≥ 40 ng/mL.

Overall, a low vitamin D status was present in 87.4% of patients with severe or critical COVID-19, but only 34.3% of patients with mild to moderate disease were vitamin D deficient. Mortality among patients with sufficient vitamin D was 2.3%, but 25.6% in vitamin D deficient patients.

“This study contributes to a continually evolving body of evidence suggesting that a patient's history of vitamin D deficiency is a predictive factor associated with poorer COVID-19 clinical disease course and mortality,” said study co-author Michael Edelstein a professor at Bar-Ilan University. “It is still unclear why certain individuals suffer severe consequences of COVID-19 infection while others don’t. Our finding adds a new dimension to solving this puzzle.”

Article 3 (Included in Study 1)

Eating Less Dairy Fat Increases Risk of Heart Disease, Study Suggests

A lower consumption of dairy fat increases risk of cardiovascular disease (CVD) as compared to high intakes, according to new research studying some of the world's biggest consumers of dairy products.

An international team of scientists studied the dairy fat consumption of 4,150 60-year-olds in Sweden – a country with one of the world's highest levels of dairy production and consumption – by measuring blood levels of a particular fatty acid that is mostly found in dairy foods. Experts then followed the cohort for an average of 16 years to observe how many had heart attacks, strokes, and other serious circulatory events, and how many of them died.

After statistically adjusting for other known cardiovascular disease related factors including age, income, lifestyle, dietary habits and other diseases, researchers found that low levels of the fatty acid – indicative of a low intake of dairy fats – increased risk of cardiovascular disease.

The team then confirmed these findings in other populations after combining the Swedish results with 17 other studies involving a total of almost 43,000 people from the US, Denmark and the UK.

"While the findings may be partly influenced by factors other than dairy fat, our study does not suggest any harm of dairy fat per se," Matti Marklund, senior researcher at the George Institute for Global Health in Sydney and joint senior author of the paper, said in a statement.

"We found that lower levels of the fatty acid increased risk of CVD. These relationships are highly interesting, but we need further studies to better understand the full health impact of dairy fats and dairy foods," he said.

Article 4 (Included in Study 1)

Just 30-90 minutes of resistance training weekly decreases risk of premature death – new research shows

There's long been evidence that moderate aerobic exercise (e.g., walking, running, or cycling) are good for your lifelong health and well-being. Research even shows us more active people also tend to live longer, healthier lives with lower rates of disease – including cancers, diabetes, and cardiovascular disease.

But what about resistance exercise – like lifting weights? While it's thought these kinds of exercises are probably also good for health and longevity, less evidence has existed showing the benefits. But a recent study now shows that 30-90 minutes of resistance training a week may decrease risk of premature death from all causes by 10%-20%.

The team of researchers from three universities in Japan conducted a meta-analysis – meaning they pooled data from 16 separate studies looking at longevity, disease related factors, and resistance exercise. This allowed them to look at tens of thousands of participants altogether.

They found that 30-90 minutes of resistance exercise per week may strongly decrease risk of dying from all causes. More strikingly, they also found regularly performing more than three hours of strength training per week actually poses an increased risk of premature death.

They also found that the optimal amount of time spent resistance training varied when it came to preventing different diseases. For example, while 40-60 minutes of strength training per week may reduce the risk of cardiovascular disease. However, resistance training was shown to have no effect on the risk of some specific types of cancer, such as bowel, kidney or pancreatic.