KADİR HAS UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF SOCIAL SCIENCES AND HUMANITIES

THE EFFECTS OF KNOWLEDGE EVOKING STYLE AND ICON ARRAYS ON PERCEIVED RISK AND WILLINGNESS

DENİZ ÖZDEMİR

MASTER OF ARTS THESIS

ISTANBUL, JUNE 2022

Deniz Özdemir Master of Arts Thesis 2022

THE EFFECTS OF KNOWLEDGE EVOKING STYLE AND ICON ARRAYS ON PERCEIVED RISK AND WILLINGNESS

DENİZ ÖZDEMİR

A thesis submitted to

the School of Graduate Studies of Kadir Has University

in partial fulfillment of the requirements for the degree of

Master of Arts in

Psychological Sciences

Istanbul, June, 2022

APPROVAL

This thesis/project titled THE EFFECTS OF KNOWLEDGE EVOKING STYLE AND ICON ARRAYS ON PERCEIVED RISK AND WILLINGNESS submitted by DENİZ ÖZDEMİR, in partial fulfillment of the requirements for the degree of Master of Arts / Master of Arts in Psychological Sciences is approved by

Asst. Prof. Dr., Ahu Gökçe (Advisor) ……………………..

Kadir Has University

Asst. Prof Dr., Aslı Erciyes ……………………..

Kadir Has University

Prof Dr., Ayşecan Boduroğlu ……………………..

Bogazici University

Prof. Dr., Mehmet Timur Aydemir

Director of the School of Graduate Studies

Date of Approval: 10/06/2022

DECLARATION ON RESEARCH ETHICS AND

PUBLISHING METHODS

I, DENIZ ÖZDEMIR; hereby declare

- that this Master of Arts Thesis that I have submitted is entirely my own work and I have cited and referenced all material and results that are not my own in accordance with the rules;
- that this Master of Arts Thesis does not contain any material from any research submitted or accepted to obtain a degree or diploma at another educational institution;
- and that I commit and undertake to follow the "Kadir Has University Academic Codes of Conduct" prepared in accordance with the "Higher Education Council Codes of Conduct".

In addition, I acknowledge that any claim of irregularity that may arise in relation to this work will result in a disciplinary action in accordance with the university legislation.

Deniz Özdemir

____

10/06/2022

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor Asst. Prof Ahu Gökçe for encouraging me to follow a novel path and find new research perspectives. I am thankful for my thesis committee members Asst. Prof. Aslı Erciyes and Prof. Ayşecan Boduroğlu for their valuable contributions and constructive feedbacks to my thesis.

I am glad to my future colleagues I met in my Master's classes, without knowing them this process would not be such an enjoyable journey. I am especially grateful for Bengi and Hande for always being there for being a true friend by all means. I would like to thank Uğur as well, for being a great friend and partner to me. Before the greatest contributors, I am also thankful for my friends İpek and Emre, without them my life would lack joy and laughter seriously. And my adorable cat Tırmık, I am grateful for his fluffy presence.

Finally, and most significantly, I am grateful for having the best family. I am thankful and I will always be thankful for their support.

THE EFFECTS OF KNOWLEDGE EVOKING STYLE AND ICON ARRAYS ON PERCEIVED RISK AND WILLINGNESS

ABSTRACT

One of the most denoted cognitive errors in terms of judgments of risk probabilities was denominator neglect, and it happens when people pay more attention to numerators (subsets) and neglect the denominators (superordinate sets) during risk judgments, and it arises from the "gist-pop-out" of the numerator according to Fuzzy Trace Theory (FTT) and is defined as a class inclusion error. Previous studies which focus on visual demonstration of risk probabilities suggest that visuals help improve risk comprehension (reduce denominator neglect) and gist/verbatim knowledge extraction. Icon arrays are known to be proper visual communicators for displaying part-to-whole relationships via showing numerator and the denominator visually and utilized in reducing denominator neglect. The main objective of the thesis is to investigate risk evaluations under different risk scenarios (Experiment 1A: non-medical; Experiment 1B: medical) and to investigate the effects of visual aids, knowledge instruction types, and numeracy, health literacy on complex risk judgments, confidence levels and willingness to choose better outcomes within the overarching scope of FTT. The main hypotheses are as follows: Visual aids and instruction types will help mitigate denominator neglect and then in turn, improve risk evaluations, and higher numeracy will yield better risk evaluation accuracies. The findings demonstrated contrary results in terms of visual aids, which have been known for mitigating the effects of denominator neglect, the effects of visual aids were not replicated. Instruction types did not have the main effect on risk evaluations to help to reduce denominator neglect. Numeracy was the most prominent factor in risk accuracy, confidence in answers, and willingness. Health literacy did not show the main effect on willingness. Denominator neglect was found only in Experiment 1B.

Keywords: Fuzzy trace theory, Denominator neglect, Base-rate neglect, Visual aids, Icon arrays, Risk judgments, Risk reduction accuracy

BİLGİ ÇAĞRIŞTIRICILARININ VE İKON DİZİLERİNİN ALGILANAN RİSKE VE İSTEKLİLİĞE ETKİSİ

ÖZET

Payda ihmali, risk yargıları konusunda en çok bilinen bilişsel hatalardan biri olup bireylerin riskleri yargılarken paya (alt kümeye) daha fazla dikkatlerini verip paydayı (üst kümeyi) ihmal ettiklerinde ortaya çıkar. Bulanık İz Teorisi'ne göre payın özet yargısının çok fazla ön plana çıkması sonucunda oluşan bu durum sınıfa katma hatası olarak tanımlanır. Risk olasılıklarının görsel olarak gösterilmesi konusunda önceki çalışmalar görsel kullanımının risk kavramada (payda ihmalini azaltmada) ve özet/detaylı yargıların çıkarımında fayda sağladığını öne sürer. İkon dizilerinin pay ve paydayı görsel olarak göstererek parçadan-bütüne ilişkileri göstermek için uygun görsel iletişimciler olduğu ve payda ihmalini azaltmada kullanıldığı bilinmektedir. Tezin temel amacı, farklı bağlamlardaki risk senaryoları altında (Deney 1A: tıbbi olmayan; Deney 1B: tıbbi) risk değerlendirmelerini araştırmak ve Bulanık İz Teorisi kapsamında görsel yardımcıların, bilgi yönerge türlerinin, numerik yeteneğin, sağlık okuryazarlığının karmaşık risk yargıları, güven seviyeleri ve daha iyi seçeneği seçme isteği üzerindeki etkilerini incelemektir. Ana hipotezler aşağıdaki gibidir: Görsel yardımlar ve talimat türleri, payda ihmalini azaltmaya yardımcı olacak ve risk değerlendirmelerini iyileştirecek; yüksek numerik yetenek fazla risk değerlendirmelerinde daha fazla doğru cevaba yol açacaktır. Bulgular, payda ihmalinin eski çalışmaların aksine görsel yardımcılar açısından zıt sonuçlar gösterdi, görsel yardımcıların risk değerlendirmeleri üzerinde etkisi ve payda ihmalini azaltmaya yardımcı olmak için verilen yönergelerin risk değerlendirmeleri üzerinde etkisi saptanamadı. Numerik yeteneğin hem risk doğruluğunda, hem de cevaplara olan güvende ve isteklilikte en belirgin faktör olduğu bulundu. Sağlık okuryazarlığı, risk azaltımı sağlayan tercihe yönelik isteklilik üzerinde temel bir etki göstermedi. Payda ihmali sadece Deney 1B'de bulundu.

Anahtar Kelimeler: Bulanık iz teorisi, Payda ihmali, Temel-oran ihmali, Görsel yardımcılar, İkon dizileri, Risk kararları, Risk azaltma doğruluğu

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONMYMS AND ABBREVIATIONS

CVD	Cardiovascular Disease
E	Estimate
FTT	Fuzzy Trace Theory
NR	Non-risk
R	Risk

1. INTRODUCTION

Perhaps the decision-making process is the most important mechanism that affects and navigates our lives. Individuals make a decision in every possible scenario, from the most basic scenarios, such as deciding to drink coffee or tea in the morning, going out or not, where or when to go out, and what to wear, to the most complex ones, such as gambling with a risk of loss of a fortune, investing money on goods (i.e., car, house, stock), and having a surgery/chemotherapy/drug or not. Eventually, some decisions and their outcomes followed by them involve some risk at varying levels for each individual, and these risk judgments may impact individuals' lives financially or health-wise dramatically.

Economic decisions are complex and a well-studied area (Park & Cho, 2019; Schurr, 1987; Tversky & Kahneman, 1973, 1981) from different theoretical approaches such as Expected Value/Utility Theory, which assumes that individuals calculate the probabilities straightforwardly and individuals are rational when they make decisions, whereas Prospect Theory claims the opposite (see for the review Reyna, 2018; Kahneman & Tversky, 1979; Schurr, 1987; Tversky & Kahneman, 1973). Affect also play a role, and individuals who were loss-aversive before can become risk-takers at the very same probabilities depending on how the gains and losses are framed (Ashby, Rakow, & Yechiam, 2017; Oliver, 2003; Reyna, Chick, Corbin, & Hsia, 2014; Reyna, 2018). Besides studies that focus on financial judgments, studies conducted in health-related fields can be subject to the same asymmetric reasoning process of health risks (Oliver, 2003; Reyna, 2012). The improvements in health decisions surely bear an applied value in improving health decisions and disease prevention communication (Broniatowski & Reyna, 2020; Reyna, 2020; Reyna, Broniatowski, & Edelson, 2021). The perceived risk of a treatable disease or a treatable disorder could alter the attributions and decisions related to the treatment (i.e., vaccine and vaccination behavior) (Reyna, 2012), adherence to the treatment, professionals' suggestions, and prevention techniques (i.e., avoiding unprotected sex against sexually transmitted diseases, taking prescribed medication in to avoid heart disease) (Hux & Naylor, 1995; Wilhelms & Reyna, 2013). The leading research interest in decision-making is how individuals deal with simple or complex decisions involving risks and benefits.

Some main mechanisms guide the decision-making process, including heuristics (Reyna, 2018; Turpin et al., 2020), risk judgments, and literacy (Fraenkel et al., 2018; Ghazal, Cokely, & Garcia-Retamero, 2014; Lipkus, Samsa, & Rimer, 2001). These factors differ due to individual differences, such as numeracy skills (Reyna & Brainerd, 2008) and health literacy (Berkman, Davis, & McCormack, 2010). In addition, some formal models build a framework upon risk judgments and decision-making processes. These models contextualize decision-making processes arising from probability problems and risk evaluations within the rational and non-rational processing axis (i.e., Expected Utility or Prospect Theory). Beyond them, Fuzzy Trace Theory (FTT) (Reyna & Brainerd, 1995) approaches the decisional operations differently by centering the mental representations rather than the process itself (i.e., Type I and II) (Thompson et al., 2021). According to FTT, individuals do not show intuitive or rational behaviors because of the processing merely. Instead, individuals both retain detailed and core levels of representations, and detailed representations are prone to fade and leave the core level representations behind. In the present thesis, the aim was to investigate the individuals' decision-making processes in risky situations within the scope of FTT in medical and non-medical contexts and how these representations claimed by FTT interact with the visual aids that help decision accuracy. The following sections will discuss FTT and its impact on decisional processes.

1.1 Fuzzy Trace Theory

Fuzzy trace theory (FTT) is one of the most studied decision-making frameworks. FTT suggests that individuals who are about to make inferences, intuitions, or decisions tend to build upon *"the fuzzy trace"* of their reasoning processes. This "fuzzy trace" comes from the information processing preference, which refers to a continuum in representation formation with two main structural elements of FTT: gist and verbatim representation of the knowledge

(Reyna & Brainerd, 1991). Gist representations (bottom-line, simplified meaning of the knowledge) and verbatim representations (exact, precise, and objective knowledge of the values) are two main foundations of FTT (Blalock & Reyna, 2016). Before the final decision or judgment, individuals can encode and store the information in two separate and parallel dimensions, which are verbatim and gist, but the information processing results in a fuzzy but more stable structure of knowledge, and this is called fuzzy processing preference (Reyna & Brainerd, 1991). In other words, people keep the *"bottom-line"* meaning of the information rather than retaining the information with precise fidelity (Reyna & Brainerd, 2008). Consider watching the news about three successfully developed vaccines (A, B, and C) against infectious disease with a 2% mortality rate and hearing the anchor-man mentioning one individual for A and B and two individuals for C showed an anaphylactic reaction after receiving the shot (A= .00001; B=.000001; and C=.000002) which yielded to serious side effects including hospitalization. The decision to have (or not to have) one of the vaccines and to pick the best option depends on the knowledge extraction from that information. Generally, people do not serially process the information from numbers or statistics. Instead, they process them in a parallel fashion to make meaningful inferences. In the previous example, the viewers can encode the gist of the risk caused by vaccination as "some" or "none" (nominal) by putting them in clusters (bins) and "greater" or "smaller" (ordinal) by comparing the vaccines; or they can encode the actual numbers of health complications and keep them as verbatim representations for a short term (Reyna, 2012; 2018; Reyna & Brust-Renck, 2020). However, even if both verbatim and gist elements are processed perfectly, fuzzy processing preference yields that the gist representations become more accessible (and verbatim, detailed representations to be faded minutes later, see Gaissmaier et al., 2012; Ruiz et al., 2013), more impactful in decision making (see for review Blalock & Reyna, 2016; Reyna & Farley, 2006). For example, a study showed that more experienced health professionals adapted more gist-like judgments and decisions about hypothetical patients (overseeing irrelevant information and admitting or discharging the patient) than their less experienced counterparts when evaluating the patient's condition (giving attention to all verbatim information and not making certain decisions such as discharging the patient) (Reyna & Lloyd, 2006). Another study conducted with adolescents showed that gist-induced training was more impactful in avoiding risky sexual acts even after 6 to 12 months than the control group (Reyna & Mills, 2014). Another study found that gist representations of the risk are associated with risk avoidance in adolescents and young adults (White, Gummerum, & Hanoch, 2015). These studies suggest that fuzzy processing preference yields verbatim representations to fade (Ruiz et al., 2013) and gist representations to stay intact and affect the decisions more than the base level (than verbatim as well). However, these bottom-line representations might yield cognitive errors as well (Blalock & Reyna, 2016; Witteman & Tollenaar, 2012).

Although gist representations are more retrievable and more stable in terms of precision than verbatim representations across time, they might also be prone to cognitive biases. These cognitive biases play a role in faulty decision-making. Blalock and Reyna (2016) mentioned the factors that can yield gist reasoning errors: information-sourced errors (imperfect or partial information), encoding sourced errors (false gist representation), memory sourced errors (faults in retrieving respective information), processing sourced errors (class inclusion of overlapping clusters). Poorly constructed gist representations arising from cognitive biases may lead to errors. For example, an incomplete gist from a health message: "wearing masks are protective in order to reduce the particle secretion, might end up with the people wearing masks to cover their mouths only, missing the nose which also contributes spreading.

Gist reasoning errors manifest themselves through probability and risk evaluations. In one of the earliest studies of FTT (Brainerd & Reyna, 1990), children (up to age 9) were asked to judge the numbers in a set with overlapping classes of elements (i.e., seven horses and three cows as the subclasses, animal as the more generative, superordinate class). Children could indicate that there are more horses than cows in the set. However, they failed to respond to whether there are more cows (sub-class) than animals (superordinate class). This error is called "class inclusion error." With more complex questions, adults demonstrate enhanced reaction times when the questions contain overlapping classes (Reyna & Brainerd, 1991; Reyna, 1991; Reyna & Brainerd, 2008). Class inclusion errors can be described as errors in processing the probability judgments when they consist of conjoint (overlapping) cases and events. These overlapping classes are invisible at first sight, and access to the sub-classes nested in the superordinate class is easier than naming the superordinate class, such as pointing cows out is easier than pointing out the animals. Furthermore, retrieving the other subset (the number of horses) was no problem. Children up to a certain age find the task difficult. When the complexity increases, adults suffer from class inclusion errors as well. That error manifests itself as the conjunction fallacies (Linda problem, highly educated individuals think that the probability of being a "feminist" bank teller is more probable than being just a bank teller only) (Wolfe & Reyna, 2010). When the comparisons of the set of classes become more complicated, the process of separating these *"nested"* classes challenges individuals, and they tend to focus on the available information (i.e., cows and horses) rather than the constructed information (i.e., animals). This challenge reveals itself by focusing more on the numerators (where subsets are represented) rather than denominators (where the superordinate set is represented). Eventually, the gist is constructed in the wrong way. FTT explains denominator neglect as a gist reasoning error. A study by Wolfe and Reyna (2010) found that class inclusion errors increase with the increment in task complexity. For adults, these nested classes might not be the farm animals but might be the probability estimates. Furthermore, these estimates impact judgments, decision making, and risk evaluation accuracy, which might be needed to make an important decision such as taking prescribed medication to avoid a heart condition to an extent, buying a car that is known to some risk of manufacturing flaw.

1.2 Denominator Neglect

One of the cognitive errors caused by interference in processing and the focus of this thesis is denominator neglect (Reyna, 2004). *Denominator neglect* is when people misjudge the numerical ratios depending on the numerator's numerical superiority over the denominator (Garcia-Retamero, Galesic, & Gigerenzer, 2010; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012; Reyna & Mills, 2014). A classical study by Denes-Raj and Epstein (1994) investigated the denominator neglect phenomenon in probability judgments. They asked participants to pick a red item from two boxes filled with blue items. The chance of picking a red item out of blue items was 1/10 and 7/100 for each box, respectively. Even though the first box appeared to be more advantageous statistically in terms of providing a greater chance for picking the target item, participants were eager to pick from the second box due to the superior numerator (seven items are more than one item, regardless of the remaining blue items). More surprisingly, participants reported that they were aware of the odds but thought they had better chances in the low-probability box (7/100) anyway. Individual differences in thinking styles (Mikušková, 2015) and initial response dependency (Szaszi et al., 2018) might be accounted for the justification.

According to the predictions of FTT, denominator neglect occurs when disjunction between classes fails, and event frequencies (numerators) are taken as an indicator of the gist of the event (Reyna, 1991; Reyna & Brainerd, 2008; Reyna, 2008; Wolfe, 1995). Neglecting the base rate (denominators) due to the saliency of the frequencies (numerators) has recently been named "gist pop-out" (Reyna & Brainerd, 2014). The numerator's pop-out might affect the perceived probabilities of the risks by overshadowing the denominator, and sometimes risks can be falsely magnified or diminished (Valerie F. Reyna & Brainerd, 2014; Zikmund-Fisher et al., 2014). For example, consider seeing a made-up ratio comparison of the risks of dying from lung cancer at approximately % 27. In addition, consider seeing a ratio that indicates the chances of death from a zeppelin accident is 1350 out of 10.000. One can say that a rational human would easily do a basic mathematical computation and demonstrate that the latter instance equals to %13.5; therefore, the chances of dying in a zeppelin accident are not higher than dying from cancer. However, this is shown to be not the case in general. Usually, individuals rate the risk to be larger than they are when it comes with a greater numerator; in this case, it would be a competition between 27 and 1350, regardless of their denominators (100 and 10000, respectively). Therefore, dying from a zeppelin accident would yield to greater perceived risk. Yamagishi (1997) showed that when mortality level is presented within a larger range and smaller frequency (in the example above, 1350/10000), the risk is perceived as greater than the actual range due to the enlarged numerator. Similarly, a study found that individuals' risk reduction estimates of a given drug (a measure that aims to define participant's perception of the drug efficacy in a non-explicit way) were underestimated when sick subjects in the treatment group were presented with a larger frequency (numerator) and overestimated when they are presented with smaller frequency/range (Okan et al., 2011). In the study, four hypothetical scenarios about drug trials were presented as sick patients in the treatment group/treatment group and sick patients in the placebo group/no-treatment (placebo) group (sick patients defined the numerator while group sizes defined the denominator) in different numbers (800-800; 800-100; 100-800; 100-100); with constant drug efficacy (80%) and constant disease rate in the absence of a drug (10%, for example, 16 would be sick out of 800 subjects are in treatment and ten would be sick out of 100 subjects in the no-treatment group). For each scenario, participants were asked to respond how many people would get sick in a group of 1000 people who took and did not take medicine, namely, "risk reduction estimate" as the outcome variable (Schwartz, Woloshin, Black, & Welch, 1997). More cases in the treatment group than in the no-treatment group yielded lower risk reduction estimates, leading to an underestimation of the drug's risk reduction (16 is more than 10). Therefore, individuals can be confused with conjunctions and judgments that merely depend on the numerator (become more visible and accessible), manifesting as gist-reasoning errors (Barbey & Sloman, 2007; Reyna, 1991; Reyna & Brainerd, 1993).

1.3 Individual and External Factors Involved in Evaluation of Risk

Individual and environmental factors play a role in risk perception of daily life and medical contexts. This section examines individual differences and external or environmental impacts on risk evaluations.

Individual differences are significant when it comes to risk perception and decisions that involve risk. Numeracy is one of the well-sounded attributes associated with improved decision-making in health-related (Reyna & Brainerd, 2008) and non-health-related judgments (Levy, Ubel, Dillard, Weir, & Fagerlin, 2014). *Numeracy* is the ability to understand arithmetic information (Peters, 2008) and comprehension of probabilities, frequencies, and odds in general which are given in numerical format (Lipkus et al., 2001;

Lipkus, 2007). Although the former definitions were about the objective measurement of numeracy, subjective numeracy is also highly related to objective numeracy (Peters et al., 2019), which are associated but separate concepts. Subjective numeracy refers to a general confidence in arithmetic abilities (Peters & Bjalkebring, 2015).

Although the factors mentioned above were related to risky decision-making regarding health decisions, health literacy also comes up as another factor. *Health literacy* is defined as the ability to gather, interpret and evaluate simple health information (Golbeck, Ahlers-Schmidt, Paschal, & Dismuke, 2005), the ability to acknowledge health outcomes (Sheridan et al., 2011), comply with better health decisions (Berkman, Davis, & McCormack, 2010; Nayak et al., 2016), better utilization of health services (Copurlar & Kartal, 2016). In addition, recent studies implied that health literacy could be extended to a more complex and statistical concept such as health numeracy, which refers to the computation and evaluation of health risks and benefits (Golbeck et al., 2005). Although health literacy might sound more behavioral construct (i.e., adherence), intuitively, it is also related to numerical or arithmetic competence, such as being able to understand the information or risk factors given by health professionals (Golbeck et al., 2005). A study found that lower health literacy predicts a reduction in vaccination willingness (Dodd et al., 2021). In addition, health literacy is also associated with a broader framework, including general literacy (Berkman et al., 2010) and general numeracy (Golbeck et al., 2005; Lau et al., 2022). Therefore, health literacy can be considered as an extensive construct that includes accurate reasoning in health-related topics and tendency and willingness toward healthy behavior, which impacts health-related judgments in terms of willingness toward healthy choices.

In addition to individual factors that affect risk perception, other factors also guide the individuals' risk perception. External (or environmental) factors might affect the perceived risk that might depend on the saliency of the mentally represented probabilities of that risk. These representations might be formed in the short term or the long term from being exposed to various resources; for instance, media presence is one of the most notorious examples. For instance, the health risk of considering drinking coffee instead of tea in the morning might probably sound non-existent at first sight. However, when a piece of additional information is mentally represented (such as cardiovascular risks related to caffeine consumption), the perception of the risk might change in the short term. On the other hand, in the long term, another example might be that readily available stereotypes can yield to overestimating the risks, such as overestimating crime rates based on racial characteristics (Quillian & Pager, 2010). The saliency of the available or previously known information alters the risk perception. Availability bias is a cognitive error, and it suggests that when individuals find it difficult to comprehend the probabilities, they rely on the basic heuristics which are affected by the external information channels (i.e., media, society) and might yield inaccurate estimation of the probabilities or actual risks due to the perceived saliency of the risk (Tversky & Kahneman, 1973). One may speculate that media coverage showing the low (but tabloid/sensational) risks as higher than the actual level could amplify the perceived risk from the availability bias perspective. Since dying from a zeppelin accident is less risky but more sensational than cancer-related death, it would be rated riskier according to availability bias (Tversky & Kahneman, 1973). However, some contrary findings favor the preference for extracting the gist from ratios over the availability created by media framing. For example, the study by Stone et al. (1994) showed that when the high risk is given in relative format (reducing risk by half) or incident rate (reducing from .003 to .006), people are equally willing to pay for a new product in order to avoid the risk. When the low risks are given in a larger range (.000003 and .000006), people are more willing to pay for the new product when shown in a relative risk format. When the risk evoking incidents are changed in subsequent experiments according to media presence, Stone and colleagues find that when actual risk and media presence are very low (amusement park accidents) and presented in incident rate format, people were less willing to pay for improvement (i.e., paying for mechanical improvements in an amusement park). When it is given in relative risk, they are more willing to pay for safer amusement (in that case, being sold on to pay more because of the reduction of the risk *"by half"*). There were no differences between presentation formats, neither for low risk -high media coverage (airplane accidents, e.g.), for high risk - low media coverage (chainsaw accidents), nor for high risk-high media coverage (burglary, e.g.). Therefore, they claimed that individuals extract the gist of the given information, regardless of their emotional availability, which is caused by the media coverage, or their availability created by their tabloid value. Although in the previous study, some of the risks were not on the agenda of most individuals daily, even if they are considered highly risky. Some incidents can be perceived as both tabloid and prevalent in terms of riskiness, such as the COVID-19 pandemic.

During the COVID-19 pandemic, emotional availability of the risk was at an immense level, conventional or social media was bombarding the society with vast amount of numbers, and since global health authorities declared the pandemic, the risk was down-to-earth. That threat came with the enhanced levels of perceived risk and worry of the COVID contraction (Attema, L'Haridon, Raude, & Seror, 2021). In order to standardize and make understandable the actual risk, the institutions or states were conveying the information to people about the risks in various formats. For example, some gave the raw numbers of active cases daily, and some gave the case ratio regarding the population (e.g., 200 cases per 100.000 people). A study conducted by Jie (2022) asked participants about the risks of visiting a country (1 billion population) that has active COVID cases (1 million). The results demonstrated that when COVID-19 cases were given in direct numbers rather than frequencies (1 in X format), it increased the overall perceived risk. Because participants generally ignored the amount of population as the denominator. Moreover, this effect persisted even though the population (1 billion) was given in the scenario or not (Jie, 2022), which is in accordance with the gist pop-out view (Reyna & Brainerd, 2008), the risk is perceived as higher than usual when the numerator is big.

These findings indicate that even if the public or individual agenda is affected by environmental factors, the demonstration of the risk is one of the strongest predictors of risk perception, especially when the information is complex. The denominator neglect occurs even if there are other confounding factors such as availability (saliency created by environment) and affect (worry) (Jie, 2022). And as discussed in the individual factors section, risk perception is not independent of individual differences such as numeracy. However, not everyone has adequate numeracy skills; therefore, visual aids are widely used to mitigate low numeracy's effects on risk evaluations.

1.4 Risk Comprehension and Visual Demonstration of the Risk

Understanding the actual risk might sometimes be complicated for individuals. Previous research on risk perception is deeply connected with visualization techniques. Generally, visual demonstrations (such as info-graphs) are considered to be the auxiliary units of risk communication (see, e.g., Fagerlin, Zikmund-Fisher, & Ubel, 2011; Zikmund-Fisher et al., 2008). There is also an applied value of using visual demonstrations for decision-making for different contexts. For example, a study conducted in the UK with Polish immigrants showed that in addition to the health risk information text, visual aids (icon array, systemic ovals) enhanced the accurate risk perception even in the case of denominator neglect (Garcia-Retamero & Dhami, 2011). For the decisions which do not involve health information, a study showed that individuals with learning disabilities demonstrated better financial decisions and increment in task engagement when visual aids were utilized during the process (Bailey, Willner, & Dymond, 2011). Therefore, visual aids can enhance the decision-making performance by improving the comprehension and accuracy in risk perception in both health-related and daily life examples for individuals who struggle with the language barrier or have cognitive difficulties.

However, there are also drawbacks when depicting the risk with visuals. For example, Fraenkel et al. (2018) conducted a study with a medical scenario about medicine that cures a hypothetical disease but also causes a risk of infection at "small" or "very small" rates of risk. They showed that participants found small (2%) and very small (0.2%) risks understandable even without visuals. Moreover, when differentiating small and very small risks, individuals with college education did not benefit from the visuals and showed no difference in terms of willingness to start the medication, worrying about the infection caused by the medicine, or believing that the benefit of the medicine is higher than the risks (Fraenkel et al., 2018). On the other hand, individuals with lower education benefited from the visuals up to a certain point in terms of comprehending the risk. When more visual analogy was added, participants' willingness and attitude towards the medication did not benefit as intended. Individuals with higher education remained unaffected across risk conditions and visual conditions. The researchers speculated that higher education benefits health literacy through exposure, previous knowledge, and confidence. Another study examined the recall of the cardiovascular risk information of different time periods, and they claimed that visual aids might not be helpful but detrimental for short-term recall of information (Ruiz et al., 2013).

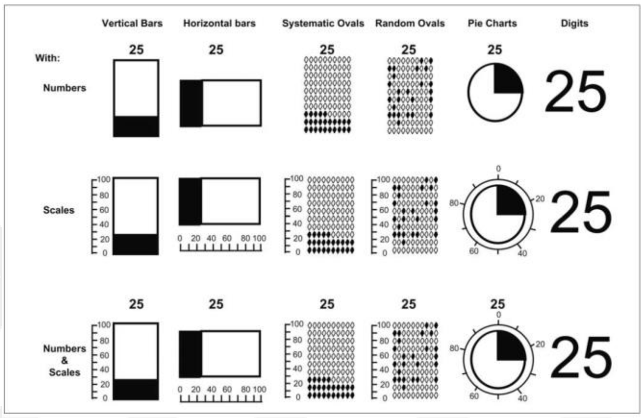
1.4.1 FTT approach to the visual formats

Briefly, the nature of the information that is supposed to be drawn from the graphical demonstrations and individual's literacy skills (i.e., numeracy, graph literacy, health literacy, general literacy) might be definitive in terms of risk communication with or without graphs. FTT offers some theoretical background regarding the type and nature of graphical demonstrations and what kind of information can be extracted from them. Health care professionals benefit from the studies that aim at improving the comprehension of the patients, and patients prefer graphical formats rather than only text (Goodyear-Smith et al., 2008), and gather more information from texts which contain both numbers and words together (Carey, Herrmann, Hall, Mansfield, & Fakes, 2018). Pictographs, tables, bar graphs, pie charts (Hawley et al., 2008; van Weert, Alblas, van Dijk, & Jansen, 2021), and even pictures that have emotional and story value (Ooms, Jansen, & Hoeks, 2020) are utilized in conveying risks and benefits of health-related topics.

Bancilhon and Ottley (2020) emphasized some critical points about decisions based on interactions between visualizations and visual reasoning. They questioned whether individuals inherently extract the information from some visualizations employing the gist or verbatim representations or whether the direct influence of visualizations yields different encoding and reasoning strategies. Studies showed that some visualizations are doing great in both evoking gist and verbatim representations, Feldman-Stewart et al. (2007) demonstrated the best option as "systemic ovals," which are also used in the clinical decision-making field as "icon arrays" or "pictographs" (See Figure 1.1). In line with that, Hawley et al. (2008) examined six graph types in terms of the efficiency at conveying the bottom line and detailed knowledge within the range of numeracy and graph literacy. In the experiment, participants were presented with a health risk scenario and two treatment options for avoiding the risk (one of the treatments has more side effects at a better success rate, and the other has fewer side effects at a lower success rate) with graphical demonstrations. Then researchers measured the retention of the knowledge with four verbatim and two gist questions, in addition to the best treatment choice (operationally defined as picking the treatment which has the higher success rate). Results showed that while "table" visualizations are better at evoking verbatim knowledge, "pie charts" are better at evoking the gist. Compatible with the benefits of using "systemic ovals," suggested by Feldman-Stewart et al. (2007), icon arrays (pictographs) are better at evoking both gist and verbatim knowledge in general.

A recent study by van Weert et al. (2021) examined participants' health knowledge derived from six different graphical demonstrations. A scenario about health risk and its treatment by providing two options is depicted with varying side effects and risks. To measure the apperceived knowledge from the graphs, researchers used multiple-choice questions, which indicated either four questions about the exact numbers (verbatim) or two questions which are core takeaway points (gist). They also measured *health literacy, numeracy,* and *graph literacy*. When the information was shown with tables, verbatim knowledge retention was greater for both age groups. Contrary to the findings of Hawley et al. (2008), tables helped the older adults extract the bottom-line meaning as well, which reflected itself as higher gist scores. Controlling for other variables, graph literacy, numeracy, and younger age were the main predictors of accurate verbatim knowledge. On the other hand, gist knowledge is best predicted by both health literacy and graph literacy. This study revealed that the tables are best for both verbatim and gist knowledge extraction, especially for older adults who can read and acknowledge the information through tables. Younger participants did not differ among graph types in terms of extracting the bottom-line meaning of the information (van Weert et al., 2021). As the authors stated, although low numeracy individuals can succeed in the tasks with at least intermediate graph literacy, being numerically literate is not the only factor in knowledge accuracy extracted from the visuals. This conclusion might lead researchers to achieve an applied solution. Although individual differences play a role in graph perception and knowledge extraction, these individual differences might not work in not all or none; rather, they can be complementary and appropriate visual demonstrations and can be developed for some chunks of individual factors in addition to the appropriate knowledge types (i.e., gist).

A. Different Visualization Formats



B. Icon arrays

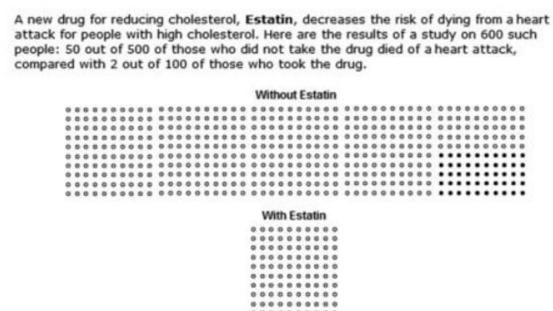


Figure 1.1 Examples of visual aids used in previous studies

Figure shows the examples of visual aids used previously by (A) Feldman-Stewart et al. (2007) and (B) Garcia-Retamero et al. (2010). Icon arrays which were used in the present study were similar to B.

1.4.2 Human – graph interaction

The studies, as mentioned earlier, have demonstrated that some associations between individual differences might play a role in visual reasoning processes through graphs. Banchilhon and Ottley (2020) asked some questions about the graphical representations and what individuals extract from them or questioned how graphical displays evoke certain knowledge types (gist or verbatim). In terms of visualization of the data, some graph types can be associated intuitively with the latter question. For example, icon arrays/pictographs that show subsets and superordinate sets might trigger gist processing (by showing the proportions visually), and tables that show the same sets of classes in numbers might trigger verbatim processing. Previous studies generally showed the associations (Bancilhon & Ottley, 2020) but lacked the explanation showing the mechanism for triggering a particular type of processing. Tables and icon arrays are both commonly used for depicting risks or proportions. By nature, tables show the numbers starkly, not the relationship between numbers. Therefore, for the table demonstrations, higher numeracy skills might be needed for making calculations and numerical comparisons through depicted numbers. Similarly, counting the dots in the icon arrays might be the common strategy for highly numerate individuals, not the part-to-whole visual reasoning. Hess et al. (2011) found that individuals with higher numeracy count the dots in the icon arrays to find the best option during risky medical decision-making. In the study, they manipulated the risk magnitude (2 or 17 out of 100 get colon cancer, e.g.) and task sequence (participants received no-trigger to count or trigger to count the colored dots in the icon array in a counterbalanced order). In no-trigger (no-count) condition, they asked: about the “general probability of the patient’s having cancer” (gist-like inference). In the trigger condition (count), they asked, “how many people are at risk like this patient?” which should be pushing participants to count the dots (verbatim-like inference). Higher or lower numeracy participants saw a higher or lower probability of risk conditions in either trigger first or no-trigger first sequences. They expected that the task sequence would not change the risk perception for higher numeracy individuals because they internally perceive the exact numbers (or, similarly, verbatim knowledge). However, lower numeracy individuals expected that triggering might result in better differentiation between low or high probability comparisons. The results partially supported their hypotheses. First, regardless of the task sequence, higher numeracy individuals correctly differentiated high or low probability. However, for low numeracy participants, triggering to count did not result in better differentiation between lower and higher probabilities. Contrarily, they were significantly better at differentiating when they were not encouraged to count first. The declining performance of the low numeracy individuals from no-counting to counting sequence overlaps with being less number-competent; therefore, they cannot compute numbers. Based on the findings of Hess and colleagues (2011), evoking a verbatim approach to interpret the data and offloading individuals with numerical calculations might not be the best option in terms of conveying health risks in health communication. Because triggering the verbatim approach can impede judgment accuracy, especially for low numerate individuals. Therefore, understanding the mechanism between the human factors (i.e., numeracy, graph literacy) and graphical designs that are specifically created to be in a way that fits best for the different individual factors and characteristics of the visual designs in order to improve risk or probability understanding.

1.5 Overview of the thesis

Previous studies provided which formats are best understood, including using words and numbers together (Carey et al., 2018) and icon arrays that can aid in health decisions (Garcia-Retamero, Galesic, & Gigerenzer, 2010; Hawley et al., 2008; Okan et al., 2011), as well as tables (van Weert et al., 2020), and which graphs are better understood depending on numeracy, graph literacy, health literacy skills, retaining the knowledge across time (Gaissmaier et al., 2012). Nonetheless, we know that graphical/numerical literacy is associated with accurate visual reasoning through icon arrays are better for accurate risk comprehension and decisions in the health domain. What we do not know is whether we can increase comprehension and compliance through eliciting gist and verbatim knowledge embedded within the messages, whether encouraging individuals to operate on specific tasks could improve their risk-related reasoning processes, and whether this influence can work within even the most complex and challenging messages which contain nested classes, in which extensively studied as the denominator neglect.

In two experiments, individuals' risk perceptions and their decisions under different levels of risk scenarios are examined in non-medical and medical contexts. Among the various options of visually aiding graphs, oval icon arrays were used in both experiments. The aim was two-folded. Although previous studies supported that visual aids help individuals to understand the risks better (Garcia-Retamero, Galesic, & Gigerenzer, 2010; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012), there are some blurry points of the addition of visual aids might be ineffective for risk understanding (Ruiz et al., 2013), especially for individuals with higher numeracy (Fraenkel et al., 2018), additive visual aids (illustrations with gist-like instructions) might be confusing for individuals with lower numeracy in terms of risk comprehension and risk-congruent judgments (Fraenkel et al., 2018). Therefore, the first question of the study was about visual aids.

> RQ1: Do visual aids help individuals understand the risks accurately even in the case of denominator neglect?

Secondly, previous studies showed that gist inference from the given information is important. It persists longer than verbatim inferences, and it is assumed to be more effective in decision-making accuracy. However, rather than an inference or extracting the knowledge, a minimal amount of research focused on evoking the specific knowledge would help improve decision-making. Therefore, the second research question focused on this novel point.

> RQ2: Do instruction types have an effect on risk accuracy, confidence levels, and willingness?

Thirdly, the following research question was about individual differences in numeracy skills and health literacy in terms of risk understanding, confidence in answers, and willingness to choose one of the options, including varying risk rates.

RQ3: Do numeracy (Experiments 1A and 1B) and health literacy (Experiment1B) play a role in risk evaluation accuracy, confidence, and willingness ratings?

Finally, this study questioned whether the risk perceptions and related decisions vary across medical and non-medical contexts. Previous studies demonstrated that base-rate neglect and denominator neglect affect decision accuracies in various medical and non-medical contexts. However, the studies were not using the same tasks (they were separate studies conducted by separate researchers); therefore, they did not encapsulate the same tasks (or methods) to examine two different contexts.

RQ4: Do medical and non-medical scenarios differ in risk evaluation accuracy?

Briefly, the current thesis aims at addressing and discussing these research questions in two experiments.

2. EXPERIMENT 1A & 1B

Two experiments were conducted to investigate the effects of scenario type (risk vs. non-risk), instruction types (no vs. gist vs. verbatim), visual aids (aided vs. non-aided), and objective numeracy (high vs. low) on risk perception, confidence in answers, and willingness. In Experiment 1A, this relationship was examined in a non-medical context, and participants were provided with hypothetical scenarios about cars with the risk of manufacturing faults. In Experiment 1B, the medical context was adapted, and participants were given hypothetical scenarios about medicine that reduces the risk of developing cardiovascular disease (CVD).

2.1 General Method

The general method of Experiment 1A and 1B were similar, with few exceptions, which were noted in the materials and procedure sections of each experiment.

To examine the effect of simple instructions with or without a visual aid on risk evaluations, single-sentence verbatim, or gist instructions that contain information about the non-adverse cases (flawed cars or heart attack cases) in the given hypothetical medical and non-medical scenarios, which varied in their risk-evoking status. Both instructions aim to disjoint the overlapping classes by emphasizing part-to-whole relationships within the scenarios. The difference between them is whether the instructions evaluate the remaining sub-group in a gist or verbatim approach. In both experiments, instruction type and scenario type were manipulated within-subject. The instruction type factor had three levels; no instruction, gist instruction, and verbatim instruction. The scenario type had two levels; risk evoking (due to the denominator neglect) and non-risk evoking. The visual aid was manipulated between subjects and had two levels; aided and non-aided. Numeracy had two levels low and high. In both experiments, a 2x2x3x2 mixed factorial design was adapted. Variances in risk estimations (raw risk scores, risk reduction estimates, risk reduction accuracy), confidence levels, and willingness were examined as the function of variables. The second experiment adapted health literacy as the between-subject factor with two levels, low and high, to examine health literacy's effects on willingness.

2.1.1 Participants

Participants were recruited from adult population in addition to Kadir Has University students. For both Experiment 1A and 1B, Qualtrics Online Survey tool is used in both presenting and collecting the data, a priori power analysis via G*Power (Faul, Erdfelder, Buchner, Lang, 2009) showed that 124 participants are needed to detect a small to medium effect size ($f = 0.25$) with 95% power of rejecting the null hypothesis, $\alpha = .05$.

2.1.2 Materials

Experiment 1A and Experiment 1B adapted similar materials that are given in the general method section. The commonalities between materials were explained in detail in the following sub-sections. The differences will be given in the latter sections, regarding the experiment.

2.1.2.1 Scenarios

There were six scenarios about hypothetical car companies (1A) and medicine preventing heart attack (1B). The scenarios differed in their numerator/denominator ratios, 800-100, 900-150, and 1200-200 for half of the scenarios; for the remaining, these numbers were reversed as 100-800, 150-900, and 200-1200. Half of the scenarios (which begin with a larger range, 800-100, e.g.) had a pattern that indicates more risk (16 and 10, e.g.). The remaining half (which begin with a smaller range, 100-800, e.g.) had a pattern that indicated less risk.

Levels with greater or lower numerator/denominator ratios determined the presented scenario's risk signaling condition. Therefore, the scenarios that begin with a higher range are also subject to denominator neglect (see Table 2.3).

2.1.2.2 Knowledge instructions

There are two types of knowledge evoking instructions, namely gist and verbatim instructions embedded in scenarios, and the no-instruction as a within-subject control variable. Knowledge instructions are determined as a single line of sentences to be given at the end of each scenario (see Table 2.4). The instructions aimed at emphasizing sub-groups that are not mentioned in the scenarios' body paragraph. The instructions are constructed as neutral as possible to avoid the confounding effect of framing (such as cars were recalled back but not withdrawn or pulled from). Since there were two different car models with two different manufacturing flaw ratios in each scenario paragraph, knowledge instructions highlighted the remaining cluster (intact cars). *Gist instruction* is a sentence that emphasizes bottom-line information (without numerical data) about the sub-groups. Gist instruction had no numerical value; it stressed the correct interpretation that can be extracted from the scenario by presenting the frequency of not being adverse (flawed car or heart attack case) for each sub-group. Verbatim instruction manipulation followed the same procedure; the remaining subclass was given numerically (extracting the adverse case sub-group from the total group). These instructions aim at inducing a thought process regarding the type of instruction upon the numerator and the denominator. In order to follow FTT's approach to class inclusion in denominator neglect, the instructions are constructed to disjoint the overlapping classes, and they aim to point to the non-adverse cases, which are also parts of the whole (Reyna, 2018; Wolfe, Dandignac, & Reyna, 2019).

2.1.2.3 Visualization

Icon arrays were generated to show exact data points. Each icon visually depicted an adverse case (i.e., broken car, heart attack), and a set of systematic icons depicted the whole volume of groups (i.e., manufacturing size, total individuals). Since there were six scenarios with the different numerator and denominator counts and quotations, six different comparative icon arrays as visuals were presented under the written scenario. Icons were oval, either shown in grey and blue data arrays or solely grey, depending on the between-subject manipulation (See Figure 2.1 and Figure 2.7). In order to keep the visual complexity similar, the icons in the non-aided condition covered the same amount of area on the page; however, they did not aim to signal participants about the sub-groups. In the non-aided condition, participants received icon arrays that visually did not depict the scenario data, and all icons were grey. In the visually aided condition, participants received systematic oval dot arrays that depict the data in the scenario in a wholesome manner for both groups (cases and whole group range). Adverse cases were shown in blue, and the remaining cases were shown in the color grey.

2.1.2.4 Objective numeracy

Objective numeracy skill was measured via 11 item numeracy scale developed by Lipkus et al. (2001). This scale includes two multiple-choice questions and nine fill-in-blank questions. Excelling at this scale requires accurately calculating and comparing various odds, probabilities, risk percentages, and fractions, in addition to converting the decimal or verbal fractions into probabilities and fractions and vice versa. The objective Numeracy Scale is found to be measuring “global numeracy” and is acceptable in terms of internal consistency (Cronbach’s $\alpha = .78$) (Lipkus et al., 2001). Obtainable maximum and minimum scores from Objective Numeracy Scale are 11 and 0, respectively. Higher scores indicate higher numeracy abilities. Blank answers are considered wrong and scored as zero (See Appendix A.4).

2.1.3 Dependent variables

2.1.3.1 Raw risk scores

The answers that were given directly to the risk reduction estimate questions were considered as the raw risk scores that the participant attained to the options. In the first experiment, the risk scores should be assigned to both options (i.e., the raw number of anticipated faulty cars in better vs. worse models). Twelve raw scores were collected from every participant, six for the better and six for the worse options (See Table 2.5).

2.1.3.2 Risk reduction estimate scores

Schwartz et al. (1997) used the equation for calculating risk reduction estimates (See for the questions Appendix A.7). It represents and requires a more detailed way of describing the risk and the risk reduction and an indicator of denominator neglect when there are overestimation and underestimations of the actual risk. Risk reduction estimate accuracy is calculated via:

$$\boldsymbol{ACCURACY} = \frac{\mathbf{Q2 - Q1}}{\mathbf{100}}$$

Only the correct answers are planned to be considered accurate (20 and 100) and coded as correct (Okan et al., 2011; Schwartz et al., 1997) (responding to these questions 40 and 120, respectively, would produce the same result, which is 80/100, however, could not be considered as accurate, otherwise excluded from the analyses). In total, six risk reduction estimate scores were calculated per each participant (for each scenario, the same equation is reapplied) (See Table 2.8).

2.1.3.3 Willingness

For Experiment 1A, there were two willingness questions for each scenario. The first willingness question measured the willingness to buy the first (better) car, and the second question measured the willingness to buy the second (worse) car. The question asked participants how much they would be likely to buy the cars separately. For Experiment 1B, there was a single willingness question that measured willingness to use the medication to avoid the CVD risk. The response was given on the 7-point Likert scale. Higher scores indicated higher willingness (1 = not willing at all, 4 = somewhat willing, 7 = quite willing).

2.1.3.4 Confidence ratings

Participants' confidence in their answers was measured via a single 7-point Likert item. Lower scores indicated lower confidence levels. Responses ranged from 1 (I do not trust at all) to 7 (I trust completely).

2.2.4 Procedure

Participants were invited to the experiment via an anonymous link which led them to the Qualtrics Survey Page. The consent page was the first page that informed participants about the duration and scope of the research. The session continued if the participant agreed to participate and terminated immediately otherwise. The next page after the consent form informed participants about the nature of the experiment. After reading and agreeing to participate in the experiment, participants are informed about the experimental procedure and instructed to carefully examine the scenarios and answer the relevant questions as accurately as possible. They are informed about different scenarios they will encounter without any further information about the subject, number, and the number of the scenarios. Participants were randomly assigned into two groups; visually aided or visually non-aided conditions (See Table 2.1). In the non-aided condition, participants saw icons that showed the two whole groups in total color grey without showing the sub-groups of adverse cases, whether they are flawed cars [1A] or heart attack cases [1B]. In the aided condition, scenarios were given with the icon arrays that show the part-to-whole relationship visually. Then one of the scenarios was shown either with gist or verbatim instruction (or without instruction), and the instructions were presented randomly for each scenario type (See Table 2.2). By that, participants saw that each instruction type counterbalanced within scenario manipulations (risk vs. non-risk). After that, participants filled out two risk reduction estimate questions. Then participants were directed to the next page, where they were asked about their confidence level in answers. After rating their confidence on a 7-point Likert scale, participants rated their willingness on the next page (for buying better and worse car models [Experiment 1A] or for taking the medication [Experiment 1B]). Then participants continued with the next scenario. This section is completed when all six scenarios are read and answered. In the end, participants completed the numeracy scale (for Experiment 1A and 1B), health literacy scale, and CVD risk evaluations (for Experiment 1B). They are thanked for their participation and provided the contact information for further questions. The experiment took approximately 20-25 minutes to complete.

Table 2. 1 Scenario Sequence

Visual Aid	Counter-balancing Conditions	No Instruction		Gist Instruction		Verbatim Instruction	
		Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk
Aided	*1st*	800-100	100-800	900-150	200-1200	1200-200	150-900
	2nd	900-150	150-900	1200-200	100-800	800-100	200-1200
	3rd	1200-200	200-1200	800-100	150-900	900-150	100-800
Non-Aided	*4th*	800-100	100-800	900-150	200-1200	1200-200	150-900
	5th	900-150	150-900	1200-200	100-800	800-100	200-1200
	6th	1200-200	200-1200	800-100	150-900	900-150	100-800

The table shows the scenarios that are shown in each participant's trial and counterbalancing. Six sequences of trials were constructed in order to achieve equally counterbalanced distribution of experimental manipulations. Each line (out of six lines) represents a trial sequence which is demonstrated to the participants in randomized order. Therefore, each participant received all of the instruction types and scenario conditions.

2.2 Experiment 1A

This experiment aimed to investigate the potential effects of simple knowledge evoking types and visual aids on risk perception accuracy and overall willingness to choose better or worse options. Specifically, whether verbatim or gist evoking instructions reduce denominator neglect and improve willingness to choose the better option (or decrease willingness to choose the worse option) was investigated. In addition, the effect of visual aid on risk evaluations was brought into question. The effect of basic knowledge instructions and visual aids on risk evaluations, confidence ratings, and willingness ratings were examined in the case of denominator neglect. The hypotheses are as follows:

The presence of icon arrays that show part-to-whole relationships will help perform better in risk accuracy questions (correct responses) and, in turn, risk reduction estimates (eventually, reduce denominator neglect). Therefore;

> H_1: Risk evaluations will show no difference when the scenarios are visually aided and show the difference when they are not visually aided.
>
> H_2: The presence of icon arrays that show part-to-whole relationships will be associated with enhanced confidence levels and willingness to buy the better option.

Willingness hypotheses:

> H_3: Willingness to buy the better option and willingness to buy the worse option will be negatively correlated.
>
> H_4: Willingness to buy the better option and total risk reduction accuracy will be positively correlated, and willingness to buy the worse option and total risk reduction accuracy will be negatively correlated.

Knowledge instructions:

> H_5: Gist instructions will increase willingness to buy for the better option and decrease willingness to buy for the worse option more than both verbatim and no-instruction conditions.

H_6: In the no-instruction condition, due to the numerator's gist pop-out, there will be more incorrect responses in risk reduction estimates; participants will select a higher value than the actual value of actual risk reduction choice when there are less faulty cars in risk indicating group (overestimating) and will select lower value than the value of actual risk reduction choice when there are more cases in non-risk indicating group (underestimating).

H_7: An interaction effect is expected to be observed in terms of instruction types and presence of icon arrays that show part-to-whole relationship. Participants will be more susceptible to instruction types.

H_8: Gist instruction will yield to more willingness to choose the better option in low numerate participants than high numerate participants compared to verbatim or no instruction conditions.

Numeracy:

H_9: Confidence ratings will be higher in high numerate participants than low numerate participants. This discrepancy will be greater with icon arrays (visual aid condition).

H_{10}: Highly numerate participants will perform better in total risk reduction estimates, and their willingness will be higher for the better option. Low numerate participants will perform poorer in total risk reduction estimates, and their willingness to choose the better option will be lower.

2.2.1 Participants

Data were collected from 151 participants; participants who responded to risk reduction estimate questions more than 1000 [more than the base rate (1000) in the risk reduction estimate questions], and did not complete the numeracy scale were excluded from the further analyses. Consequentially 115 of them were eligible for the analyses (female N=86, male N=29). The average age of the participants was 24.2 (SD = 5.58), ranging from 19 to 59 years. All participants were either continuing their college education or had had their studies finished: 59.1% of the participants were high-school graduates (N=69), 1.7% were two-year graduates (N=2), 32.2% had their bachelor's degree (N=37), 7% of them were postgraduates (N=8). Although more than half of them were either psychology students or graduates, participants varied regarding their disciplines.

Table 2. 2 Scenario Sequence and Counterbalancing in Experiments

No Instruction		**Gist Instruction**		**Verbatim Instruction**	
A_1	**D_1**	B_2	F_2	C_3	E_3
B_1	**E_1**	C_2	D_2	A_3	F_3
C_1	**F_1**	A_2	E_2	B_3	D_3

The table shows that the counterbalancing used in both experiments. Duals in no –instruction condition were kept constant across three instruction types (No Instruction: 1, Gist Instruction: 2, Verbatim Instruction: 3), and two scenario conditions: Risk: A (800-100), B (900-150), C (1200-200); Non-Risk: D (100-800), E (150-900), F (200-1200). This design was shown in both aided and non-aided conditions. All participants received each instruction type with varying risk evoking conditions.

Table 2. 3 Distribution of Numerator and Denominators in Experiment

	Better Car		Worse Car	
Scenarios	Faulty	Production Size	Faulty	Production Size
Risk (800-100)	16	800	10	100
Non-risk (100-800)	2	100	80	800
Risk (900-150)	18	900	15	150
Non-risk (150-900)	3	150	90	900
Risk (1200-200)	24	1200	20	200
Non-risk (200-1200)	4	200	120	1200

Proportions of numerators and denominators are displayed in the table (Table 2.3). Scenarios revolved around these numbers and the condition which implied greater number of faulty cars for the better car numerically is defined as risk evoking scenario (i.e., 16>10 risk; 2<80 non-risk).

Table 2. 4 Instructions used in Experiment 1A

Instruction Type	Scenario Type	Sentence
Gist	For both risk and non-risk	Bu durumda (X / M / B / T / G/ Z) serisinden bir arabanın arızalanma oranı, (Y / M / D / L / N / P) serisindeki bir arabanın arızalanma ve üretici tarafından geri çekilme sıklığından düşüktür.
Verbatim	Risk	Bu durumda 800 adet üretilen X serisinden geri çekilmeyen araba sayısı 784, 100 adet üretilen Y serisinden geri çekilmeyen araç sayısı 90'dır.
		Bu durumda 900 adet üretilen M serisinden geri çekilmeyen araba sayısı 784, 150 adet üretilen K serisinden geri çekilmeyen araç sayısı 135'tir.
		Bu durumda 1200 adet üretilen B serisinden geri çekilmeyen araba sayısı 1176, 200 adet üretilen D serisinden geri çekilmeyen araç sayısı 180'dir.
	Non-Risk	Bu durumda 100 adet üretilen T serisinden geri çekilmeyen araba sayısı 98, 800 adet üretilen L serisinden geri çekilmeyen araç sayısı 720'dir.
		Bu durumda 150 adet üretilen G serisinden geri çekilmeyen araba sayısı 98, 900 adet üretilen N serisinden geri çekilmeyen araç sayısı 810'dur.
		Bu durumda 200 adet üretilen Z serisinden geri çekilmeyen araba sayısı 196, 1200 adet üretilen P serisinden geri çekilmeyen araç sayısı 1080'dir.

Table 2.4 shows the instructions that are used in Experiment 1A in each scenario type. All participants saw these instructions under the relevant scenario (risk vs. non-risk).

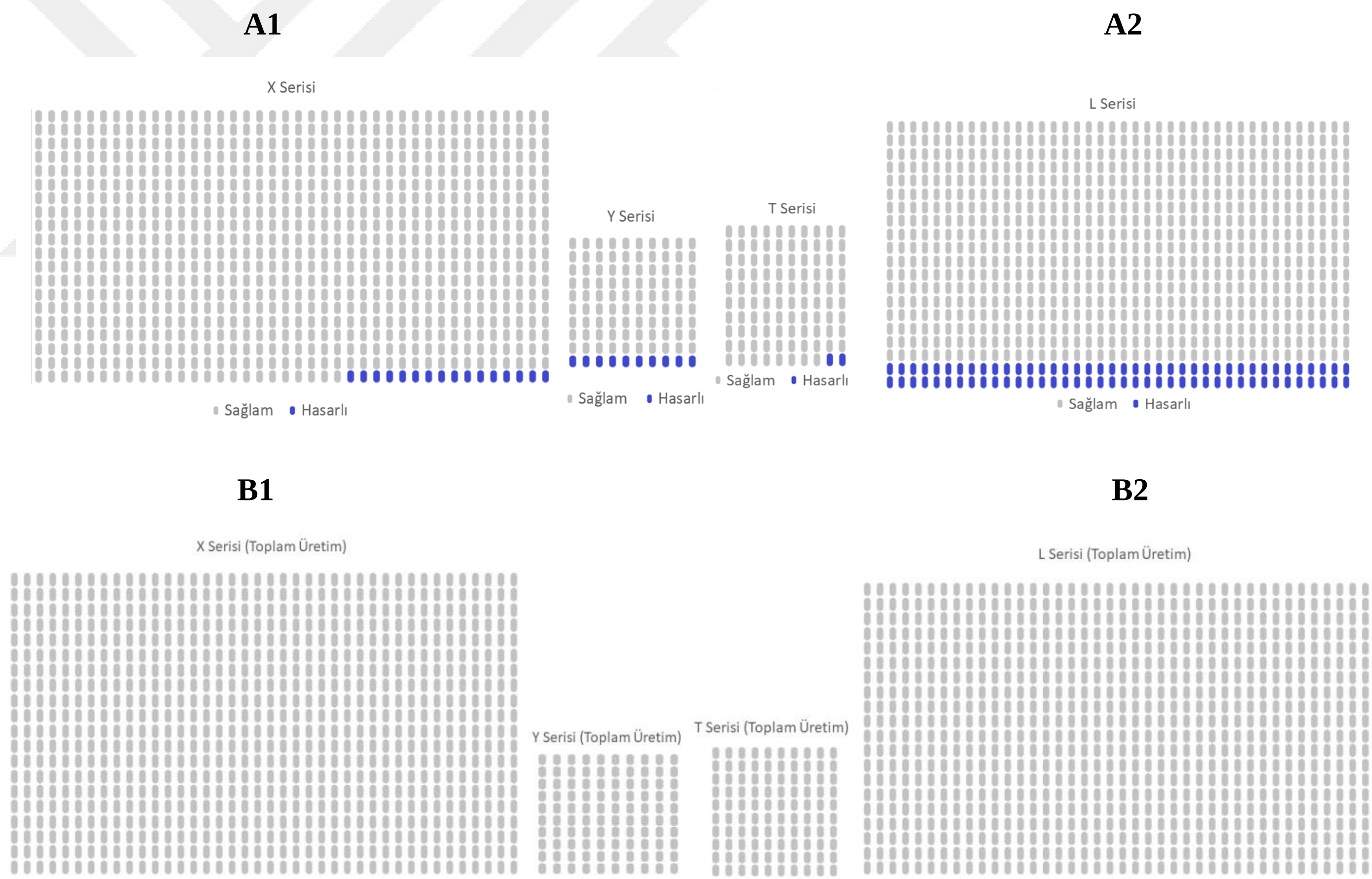


Figure 2. 1 Visuals used in Experiment 1A

The figure shows the visual demonstrations used in Experiment 1A for aided (A1, A2) and non-aided (B1, B2) with risk evoking (A1, B1) and non-risk evoking (A2, B2) scenario conditions.

2.2.2 Procedure

Participants were randomly assigned into two groups. In the visually aided condition, the participants saw the icon arrays showing sub-groups explicitly, such as running cars in the color grey and flawed cars in blue. In the visually non-aided condition, participants saw the icon arrays, which showed the two whole groups in total color grey without showing the sub-groups of running and flawed cars). There were six scenarios about hypothetical car companies producing two new car models with varying manufacturing flaw probabilities. In the scenarios, the first car option is held constant as better, which has a 2% chance of being broken down. Similarly, the second car option is held constant as worse, with a 10% chance of being broken down. The manufacturing flaw difference between the two car models is held constant at 80%. (See Table 2.1) Since there were two different car models with two different manufacturing flaw ratios in each scenario paragraph, knowledge instructions highlighted the remaining cluster (intact cars). After that, participants are shown two risk reduction estimate questions to be responded to in open-ended boxes on the same page (the text entry box is visible on the page). They were asked to indicate the expected manufacturing flaws out of 1000 for each car. Then participants rated their confidence level in answers and their willingness to buy on the next page for better and worse car models. Then participants continued with the next scenario. This section is completed when six scenarios are read and answered. In the end, participants completed the numeracy scale. They are thanked for their participation and provided the contact information for further questions. The experiment took approximately 20 minutes to complete (see Figure 2.2).

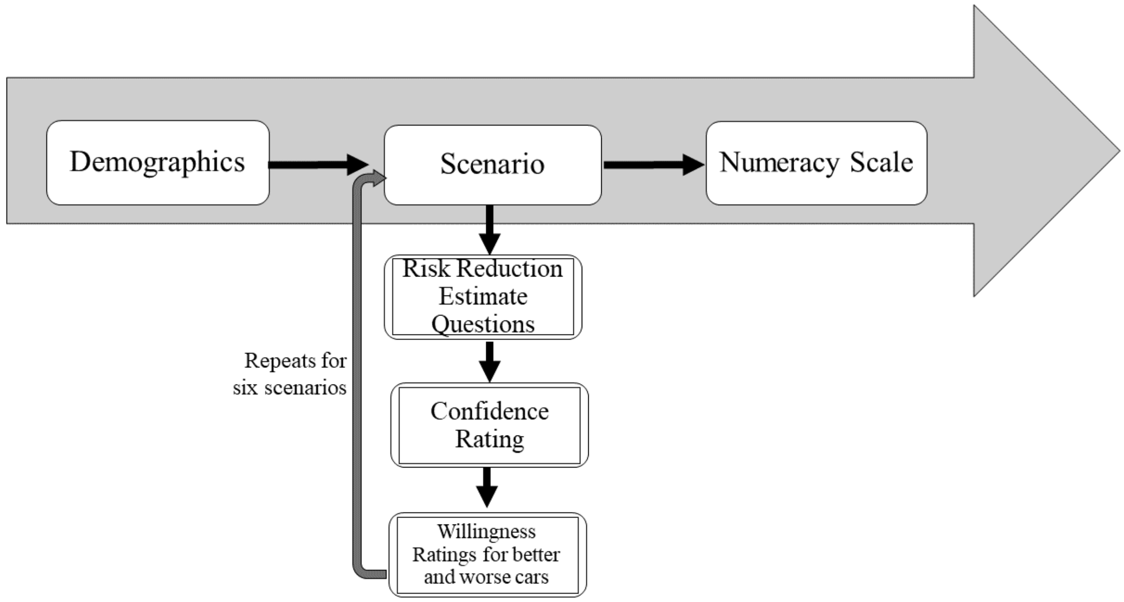


Figure 2. 2 Display Flow of Experiment 1A.

Each box refers to a separate page. After filling the demographics form (age, dominant hand, education level, branch of education), an information screen about scenarios was presented and participants are asked to skip the page when they are ready. Scenarios were presented in random order without time constraints. The page is skipped, then in the following page participants filled two risk reduction estimate questions (open ended responses, for both the better and the worse options), then rated their confidence on a single item scale (on a slider), then rated their willingness for the better and the worse options separately, (on a slider). This process repeated until all six scenarios are completed. Then participants filled numeracy scale (Lipkus et al., 2001).

2.2.3 Data coding and analyses

2x2x2x3 mixed factorial ANOVA is conducted to see the effects of the presence of icon arrays as the between-subject variable, scenario type (greater amount of faulty car cases in the better group vs. the greater amount of faulty car cases in the worse group), and instruction type (no, gist, and verbatim) as the within-subject factors, on risk evaluations (raw risk scores and risk reduction estimates), confidence levels and willingness to buy the worse and the better option. Subsequent analyses included numeracy as a between-subject factor (with two levels as low and high, achieved by median split). In the present experiment, the median value was 9 for numeracy, and the median split was applied for further analyses; participants who got scores higher than nine were marked as high in numeracy (*N*=77), and lower than nine were marked as low in numeracy (*N*=38).

Visual aid manipulation (aided vs. non-aided) and numeracy (low vs. high) were the between-subject variables. There were two within-subject factors as follows: the instruction type, which has three levels: no instruction (control), gist instruction, and verbatim instruction. The second within-subject factor was scenario type, which has two levels: risk evoking and non-risk evoking scenarios. Dependent variables were defined as the raw responses to the risk estimate questions, risk accuracy (calculated via risk reduction estimates), and confidence and willingness.

Data coding and calculations were conducted in IBM SPSS Statistics 24, and analyses were performed with Jamovi version 2.2 (The jamovi project, 2021). Single item scales (willingness to buy each one of the cars and confidence in answers) were taken as scale variables. The procedure of Schwartz et al. (1997) and Okan et al. (2011) is adapted for the risk reduction estimate scores in order to examine the denominator neglect effect (whether the estimations are correct, overestimated, or underestimated due to the experimental manipulations). Additionally, this approach was used to calculate the total correct scores of the participants for the descriptive analyses. Total correct answers were calculated regarding two responses for each scenario. Since each scenario had two risk estimate questions for two better and worse cars (one for the better and one for the worse option), total correct is only achieved when both questions are answered accurately. Therefore 6 was the highest score to be achieved for the total correct score. Besides that, raw responses (raw scorings of the participants) to the risk reduction estimate questions were also used in the analyses for better and worse cars. They are considered as the raw risk scores collected for both better and worse options in each scenario. In total, 12 answers were collected; 6 for the better (E1) and 6 for the worse options (E2) (See Table 2.5). Additionally, for each participant, six risk reduction estimate calculations were put into the analysis with three instruction types and two scenario types to see the denominator neglect effect and the distribution of their risk accuracy (i.e., correct, underestimation and overestimation).

Table 2. 5 Calculation of risk evaluations in Experiment 1A

	No Instruction		Gist Instruction		Verbatim Instruction	
RR Questions	Q1	Q2	Q1	Q2	Q1	Q2
	E_1= 20	E_2= 100	E_1= 20	E_2= 100	E_1= 20	E_2= 100
Outcome: Correct estimation of better car's risk	*=.80.*		*=.80*		*=.80*	
RR Questions	Q1	Q2	Q1	Q2	Q1	Q2
Possible Estimations (Raw Risk = E)	$E_1 \geq 20$	$E_2 < 100$	$E_1 \geq 20$	$E_2 < 100$	$E_1 \geq 20$	$E_2 < 100$
Outcome: Overestimation of the better car's risk	*<.80*		*<.80*		*<.80*	
RR Questions	Q1	Q2	Q1	Q2	Q1	Q2
Possible Estimations (Raw Risk = E)	$E_1 \leq 20$	$E_2 > 100$	$E_1 \leq 20$	$E_2 > 100$	$E_1 \leq 20$	$E_2 > 100$
Outcome: Underestimation of the better car's risk	*>.80*		*>.80*		*>.80*	

The table shows the calculations of risk evaluations. The correct answers were the same for each scenario (whether it is risk or non-risk). Risk reduction estimate questions were analyzed in three terms in the first experiment. Firstly, raw risk scores were taken into account. For example, the participant rated the first car's (the better option) risk as 20 out of 1000 and the second car's (the worse option) as 100 out of 1000. In that case, 20 and 100 were considered as raw risk scores for the perceived risk of the cars, respectively (E1 for the better, E2 for the worse). Secondly, in order to see whether there is an effect of denominator neglect or not (the risk estimates ratio was supposed to be significantly higher (and higher than .80) in risk evoking scenario conditions), Schwartz and colleagues' procedure was followed (1997). Thirdly, to see the answers' distribution, answers were coded as correct, underestimated, and overestimated (Okan et al., 2011).

2.2.4 Results

Firstly, descriptive analyses were conducted. Overall, participants tend to have high numeracy scores, ($M = 8.77$, $SD = 2.40$). Spearman's rank correlation demonstrated that numeracy score and education levels are moderately associated r (113) = .20, p = .02. Participants generally answered the risk estimate questions correctly (M= 4.44, SD =1.90). Numeracy scores and total correct answers were found to be positively and strongly correlated r (113) = .39, p <.001.

2.2.4.1 Raw risk evaluations for the better option

Participants' overall evaluation for the better car was examined. The better choice was always the first option which is represented as risk or no-risk scenarios. Also, these questions are given in three instruction types as no instruction (control), gist instruction, and verbatim instruction. Mauchly's sphericity test indicated that the assumption of sphericity has been violated for instruction type (W=.43, p <.001), therefore Greenhouse-Geisser correction is applied and degrees of freedom are corrected (ε =.636) (Navarro & Foxcroft, 2019, p.355).

The results showed that the within-subjects factors have no main effect on participants' risk ratings. There is no main effect of instruction type on risk ratings of participants, neither for no-instruction (M 28, SE = 2.69, 95% CI [22.7, 33.3]), nor for gist instruction (M = 34.8, SE = 5.45, 95% CI [24.0, 45.6]), and nor for verbatim instruction (M = 28, SE = 2.75, 95% CI [25.5 , 35.5]), F (1.27, 142.38) = 1.34, p=.256, $\eta^2 p$=.012. Scenario type has no main effect of participants' risk ratings. Neither for risk evoking scenarios (M = 29.7, SE = 2.84, 95% CI [24.1, 35.3]), nor for non-risk evoking scenarios (M = 28, SE = 3.72, 95% CI [23.5, 38.2]) participants did not differ significantly across two scenario conditions F (1, 112) = 0.9, p=.764, $\eta^2 p$=.001.

The main effect of visual aid as the between subject factor was not significant. Visual aids have no impact on participants' evaluations of the better cars. Results showed that participants' risk scores did not differ across two conditions, whether there is a visual aid (*M* = 32.4, *SE* = 3.71, 95% CI [25.1, 39.8]) or not (*M* = 28.2, *SE* = 3.84, 95% CI [20.5, 35.8]), F (1, 112) = 0.64, p = .427, η^2p=.006.

The results also indicated that, there was a two-way interaction between scenario type and visual aids F (1, 110) = 6.33, p = .013, η^2p=.054. This result indicated that, for the better option when there is no visual aid, the perceived risk tends to be overestimated (digressing from the correct answer) in risk evoking scenario whereas non-risk evoking scenarios tended to be perceived as less risky (regressing towards the correct answer, less overestimation) (See Figure 2.3).

2.2.4.2 Raw risk evaluations for the worse option

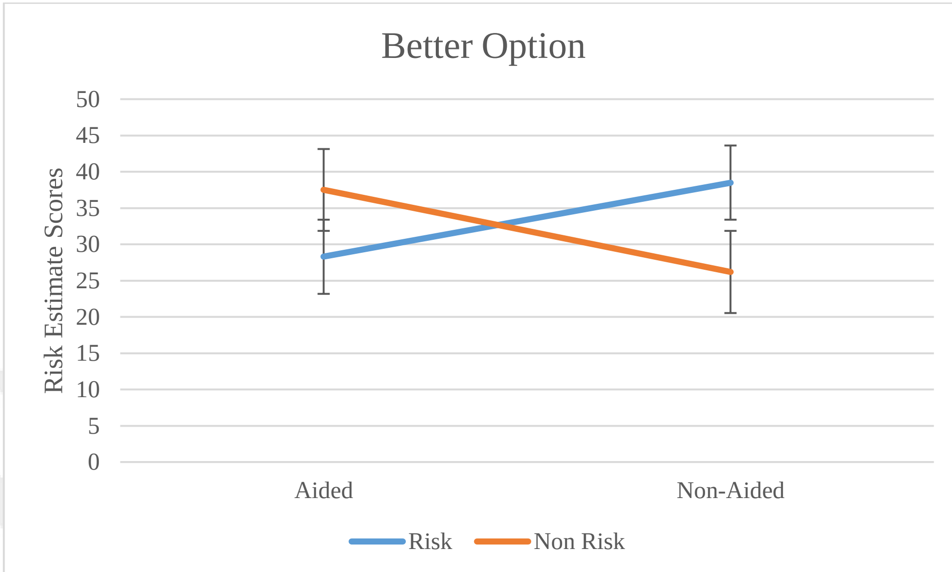
Similar analysis was conducted for the worse option as well. Mauchly's sphericity test indicated that assumption of sphericity has been violated (W=.92, p <.001), therefore Huyhn-Feldt correction is applied, and degrees of freedom are corrected (ε =.940) (Navarro & Foxcroft, 2019, p.355). Consistent with the better option, within subject manipulations were not statistically significant. Instruction type has no main effect on risk scorings neither for no instruction (*M* = 106, *SE* = 4.27, 95% CI [97.2, 114]), gist instruction (*M* = 122, *SE* = 10.27, 95% CI [101.1, 142]), nor for verbatim instruction (*M* = 103, *SE* = 7.23, 95% CI [89.1, 118]), F (1.88, 208.71) = 1.87, p = .159, η^2p=.017. Scenario's main effect was also statistically non-significant both for risk evoking conditions (*M* = 104, *SE* =4.42, 95% CI [95.3, 113]) and non-risk evoking conditions (*M* = 116, *SE* = 7.94, 95% CI [100.6, 132]), F (1, 111) = 2.77, p = .143, η^2p=.019. Visual aids as the between subject factor, have no effect in participants' evaluations of the worse options, even when the scenario is visually aided (*M* = 115, *SE* = 6.79, 95% CI [101.1, 128]), or not visually aided (*M* = 106, *SE* = 7.10, 95% CI [91.8, 120]), F (1, 111) = .79, p = .377, η^2p=.007.

The results also indicated that, there was a two-way interaction between scenario type and visual aids F (1, 109) = 4.33, p = .040, η^2p=.038. This result indicated that, for the worse option, when there is no visual aid, non-risk evoking scenario tends to be less overestimated by participants (approaching towards the correct answer) (See Figure 2.3).

2.2.4.3 Risk accuracy and inaccuracy

Participants' responses were coded as accurate (when the risk reduction is equal to =.80, coded as 1, and only assumed as correct when both responses two of the risk reduction estimate questions are correct, i.e., 20 and 100 respectively; not when 40-120), and inaccurate [when the risk reduction is smaller than .80, underestimation (coded as 2); when the risk reduction is greater than .80, overestimation (coded as 3) for better car's comparative risk evaluation]. Scenario type, instruction type, and visual manipulation failed to reach a main effect on the accuracy of the responses, (Figure 2.4). Only numeracy has main effect on accuracy of the risk estimations $F\ (1, 108) = 13.37$, $p = <.001$, $\eta^2p=.110$. Individuals with higher numeracy had more correct answers compared to the individuals with lower numeracy, and eventually were better at not estimating the risks under or over the actual level (See Figure 2.5). The only two-way interaction that reached the significance was between instruction type and the scenario $F\ (2, 216) = 4.22$, $p = .016$, $\eta^2p=.038$.

A. Risk Evaluations of the Better Car



B. Risk Evaluations of the Worse Car

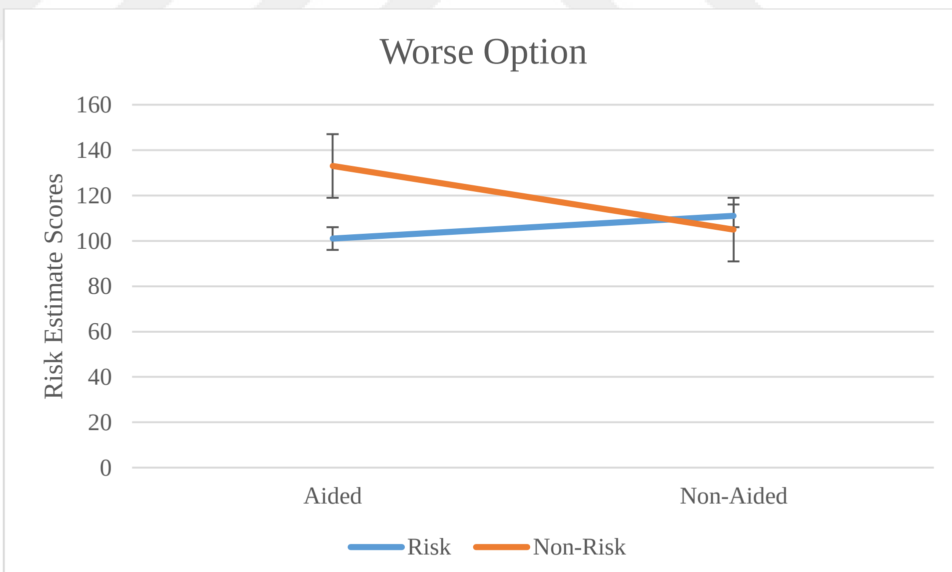


Figure 2. 3 Risk Evaluations of Better Car

The graphs show the risk evaluations of better car (A) and worse car (B) the function of scenario type and visual aids. Y-axis refers to the raw responses that are given to risk reduction estimate questions. Error bars represent one standard error.

A. Proportion of risk estimations in risk and non-risk scenarios

B. Proportion of risk estimations in visual aid conditions

C. Proportion of risk estimations across instruction types

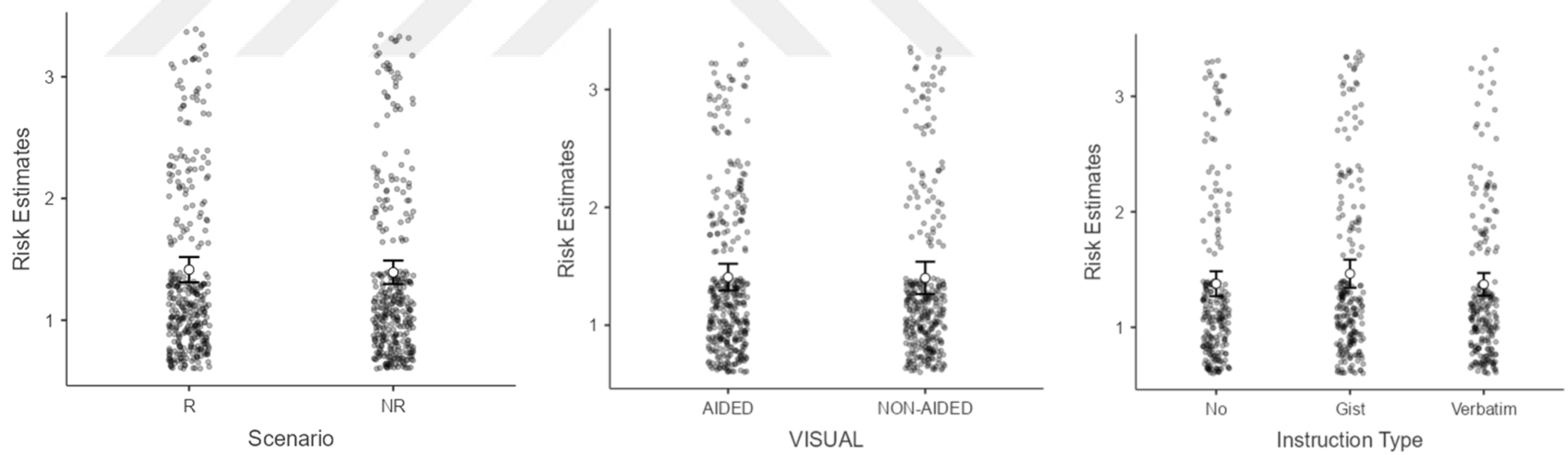
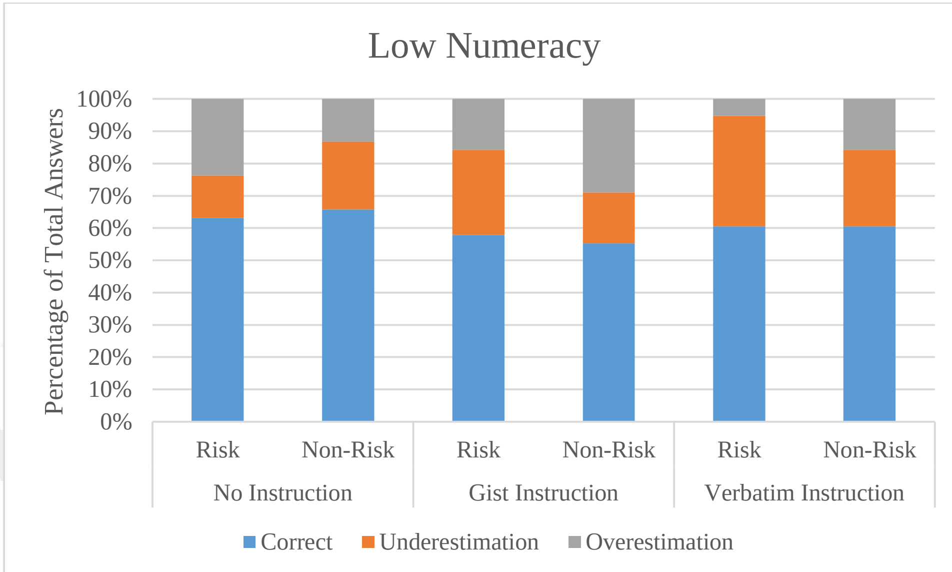


Figure 2. 4 Distributions of Correct and Incorrect Responses in Experiment 1A

First graph (A) shows participants' correct, and comparative risk estimates (underestimation and overestimation of the risk) of better and worse cars together when the scenario indicated risk (R) or not (NR). Second graph (on the middle) shows the same risk estimates in visually aided or non-aided conditions. Third graph (on the right) shows the risk estimates when the scenarios are given in three different instruction formats (no, gist, verbatim). Dots represented observed scores of the participants. Numbers on Y-axis (i.e., 1, 2, 3) represented correct (1), comparative overestimation of the better car's risk (2), and comparative underestimation of the better car's risk (3). Dots represented observed scores of the participants. Error bars represented one standard error. Accumulation of dots in 1 shows the frequency of the responses that are correct, in 2 that are incorrect and overestimated, in 3 that are underestimated.

A. Proportion of Correct and Incorrect Answers in 1A (Low Numeracy)



A. Proportions of Correct and Incorrect Answers in 1A (High Numeracy)

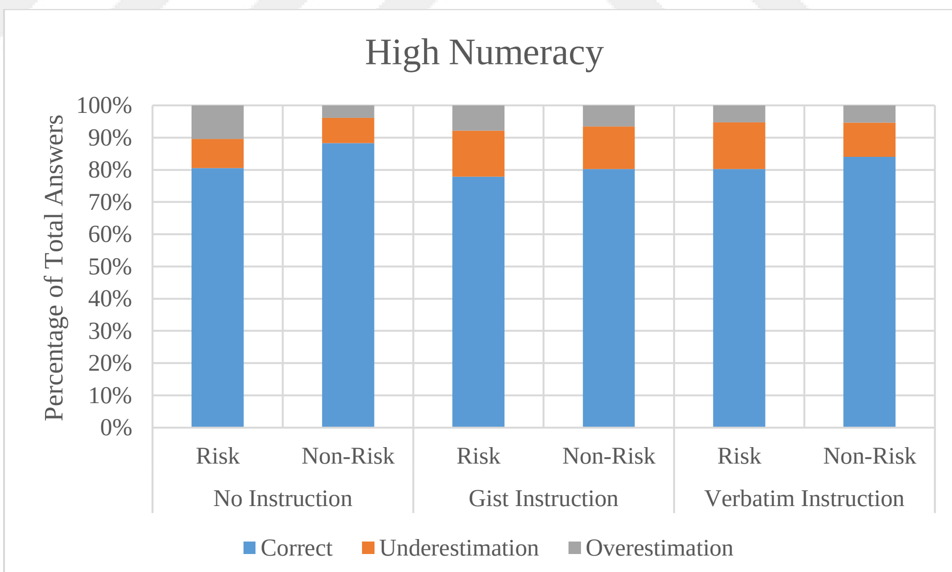


Figure 2. 5 Percentages of Correct and Incorrect Answers in 1A

The graph shows the percentages of participants whose answers are accurate, or inaccurate (comparative overestimation and underestimation of the better car's risk status) as the function of scenario and instruction type manipulations for participants with both low (A), and high (B) numeracy.

2.2.4.4 Confidence ratings

Numeracy was one and only significant predictor for confidence ratings $F(1, 110) = 25.58$, $p < .001$, η^2p=.189. Individuals with less numeracy were less confident about their responses to risk reduction estimate questions (*M* = 4.58, *SE* = .22, 95% CI [4.14, 5.01]) than individuals with high numeracy (*M* = 5.92, *SE* = .15, 95% CI [5.62, 6.21]) (See Figure 2.6A).

Main effects of scenario type, instruction type, and visual manipulations failed to reach statistical significance. There was a significant two-way interaction between instruction type and visual manipulations on confidence levels $F(2, 220) = 4.15$, $p = .017$, η^2p=.036. Three-way interaction between scenario, visual, and numeracy was also statistically significant $F(2, 220) = 10.614$, $p < .001$, η^2p=.088. None of the remaining interactions were significant (See Figure 2.6B).

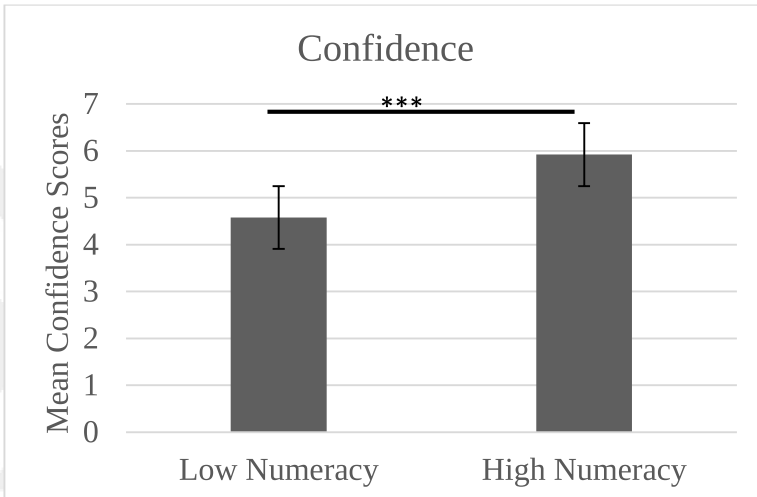
2.2.4.5 Willingness

The correlation between willingness to buy the better and worse cars was not significant $r(115) = .054$, $p = 717$ as expected. Willingness to buy the better option was significantly correlated with total correct scores $r(115) = .29$, $p = 002$.

Analyses were conducted to examine the willingness to buy the better car. Mauchly's sphericity test showed that assumption of sphericity has been violated for instruction type (W=.94, p =.027), therefore Huynh-Feldt correction is applied and degrees of freedom are corrected (ε =.956) (Navarro & Foxcroft, 2019, p.355). None of the within subject factors (instruction types, scenario types) were statistically significant. Numeracy and using visual aids had no main effect on willingness to buy the better car neither. In addition, none of the interactions were statistically significant.

Overall, instruction type, scenario type or using visual aids had no significant impact on willingness to buy the worse option. However, willingness to buy the worse option was affected by the numeracy significantly (See Figure 2.7). Participants low in numeracy rated their willingness to buy the worse car higher (M = 2.98, SE = 0.18, 95% CI [2.63, 3.34]) than their counterparts who are high in numeracy (M = 2.47, SE = 0.12, 95% CI [2.23, 2.71]), F (1, 111) = 5.64, p = .019, $\eta^2 p$ =.048.

A. Confidence Ratings



B. Interaction (Visual*Scenario*Numeracy)

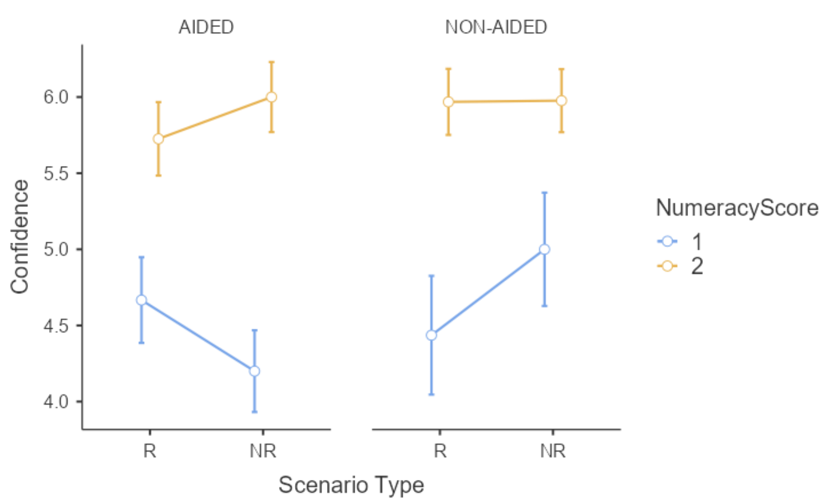


Figure 2. 6 Confidence Ratings of Participants in Experiment 1A

The graph (A) shows the confidence in responses between individuals with low and high numeracy (***p <.001). Error bars stand for one standard error. Graph (B) shows the significant interaction (p<.05) between visual aids, scenario types (R=risk; NR=non-risk) and numeracy scores (1=low, 2=high).

Figure 2. 7 Willingness Ratings of Participants in Experiment 1A

Willingness ratings for better option and the worse option among participants with high and low numeracy. Error bars stand for one standard error. (*p<.05)

2.1.5 Discussion

Visual aids had no main effect on risk evaluations and total correct answers; therefore, the first hypothesis was not supported. Risk evaluations showed no difference when the scenarios were visually aided or not. Secondly, visual aids were not found to be associated with confidence levels nor the willingness to buy the better option. Therefore, the second hypothesis was not supported. This situation can be explained in two terms. Firstly, visual manipulations might not have worked as it is intended. Since participants saw and skipped the page with the scenario and visual aids, they might not have been attentive to the visuals (this will be further discussed in the limitations section). Secondly, the sample was highly educated and high in numeracy; there is supportive evidence toward high numeracy and highly educated individuals who are good at evaluating the risk whether they are given with numbers and numbers + visuals (Fraenkel et al., 2018). Another explanation is that high numeracy individuals' icon array perception might be due to the differences in information processing driven by oval icon arrays (Hess, Vischers & Siegrist, 2011). Individuals with higher numeracy might be counting icon arrays to cross-check, which is also shown not to be the optimum option for individuals with low numeracy but preferred by high numeracy individuals when faced with icon arrays (Hess, Vischers & Siegrist, 2011; Zikmund-Fisher et al., 2014).

Contrary to the willingness hypotheses, willingness to buy better and worse options was not significantly negatively correlated; therefore, the third hypothesis was not supported. The second willingness hypothesis was partially supported. Total correct risk estimates and willingness to buy the better car were correlated. However, there was no significant correlation between willingness to buy the worse car and total risk estimate accuracy. This can be interpreted as participants with higher risk accuracy also tend to make better decisions. On the other hand, an exploratory analysis showed that willingness to buy the worse option was not negatively correlated with total risk estimate accuracy. This finding can be interpreted as individuals who are aware of the risks of two options (2% and 10%) but did not differ in their willingness. This effect was observed for very small risks before, although participants were aware of the risks (i.e., the side effect of a medication); in that study, attitudes and willingness to use medication were not affected by the risk ratios (Fraenkel et al., 2018). In addition, numerical incompetency showed a significant effect on willingness decisions. Participants who are lower in numeracy rated their willingness to buy the worse option more than their high numeracy counterparts. This overlaps with the finding that lower numeracy is associated with worse financial decisions (Ghazal, Cokely, & Garcia-Retamero, 2014; Peters et al., 2019) and higher numeracy with better decisions (Levy et al., 2014).

When risk evaluations for the separate options (better car and worse car), visual aids results yielded an interesting pattern. Better car suffers from the overestimation of the risk in risk evoking scenarios (in which the better car was given in larger group size and therefore larger numerator) from being non-aided. On the other hand, the worse car suffers from the overestimation of the risk in non-risk scenarios (in which the worse car was given in larger group size and therefore larger numerator) from being aided. This might be due to the buffering effect of visual aid with the denominator neglect. In the non-aided condition, scenarios that are risk-evoking (the group size, therefore the numerator is greater in the better option). Therefore, the risk-evoking scenario type, which is left alone without visual aid, might be overestimated, and this finding is in accordance with the literature (Okan et al., 2011). However, in aided condition, scenarios that are non-risk evoking (the group size therefore the numerator is greater in the worse option) show more blue icons in the worse option compared to the better option. Therefore, the risk demonstrated by the scenario type and visuals might cause an additive effect which might yield to overestimation of the worse car's risk. This is an interesting finding that, when comparing two options, individuals might overestimate the risks of the worse options when they are visually aided, but not the better options. When the risk evoking scenario is already exacerbating the perceived risk compared to non-risk evoking scenario (due to the denominator neglect), the visual aids might actually be helpful by showing the part-to-whole proportions by making it more interpretable for the participants for the better option. However, when the non-risk evoking scenario is supported with visual aids, this might cause to exacerbate the perceived risk for the worse option due to gist-pop out of the icon arrays that stand for worse option. Or, since the denominator neglect is not observed (scenario type had no effect on risk evaluations), inflated risk of the worse car might indicate that, in comparative risk evaluations, individuals might be even more risk aversive towards the worse options because of the numerator's superiority depicted by icons (flawed cars 2 vs. 90) (Ancker, Senathirajah, Kukafka, & Starren, 2006). This leads us to the discussions around the "transparency" of the visual aids themselves in terms of the type of information conveyed. In this experiment, although icon arrays were shown to be good at demonstrating the part-to-whole relationships, they can also be questioned for "false alarm" rates for the sub-group it represents (Kurz-Milcke, Gigerenzer, & Martignon, 2008). Briefly, visual aids might be misleading for risk evaluation accuracy when it comes to comparative decisions.

Hypotheses about the effects of knowledge instructions were not supported by the present findings. There was no main effect of instruction types on any dependent variables. There was a significant interaction between scenario type and instructions and can be interpreted with an advantage of the no-instruction condition's in terms of risk evaluation accuracy when the scenario is not risk evoking. Besides that, no other instruction type had an effect on risk perception accuracy. Willingness is also not affected by instruction types. Therefore, none of the hypotheses about instruction types were supported.

A significant interaction between visual and instruction types on confidence levels indicated a reverse pattern; in aided condition, confidence levels were lower in non-aided condition when the scenario is not instructed and instructed in a verbatim fashion, and higher in the gist. In the non-aided condition, confidence levels were higher in non-aided condition with no and verbatim instructions and lower in gist instruction (given that none of the post-hoc comparisons were significant). Although information recall and inference are widely studied topics, knowledge evoking is one of the novel concepts regarding risk evaluations in terms of FTT's gist and verbatim approach. Previous findings on knowledge evoking suggested that gist evoking is an effective method to decrease risk-taking behavior (Cho, You, & Choi, 2018; Reyna & Mills, 2014; White et al., 2015) and more effective in terms of message conveying and spread of the (mis)information (Reyna, 2020), the present results did not show favor for evoking gist information over the other information types. Although the study by Wolfe and Reyna (2010) took a more extensive way of separating overlapping classes via a step-by-step reasoning process, the questions used in the study did not differentiate between gist and verbatim reasoning processes as in the present study. Another study that focuses on the effects of knowledge types (gist and verbatim) on framing bias suggested that fine-tuning the risk evaluations can be achieved via verbatim paraphrasing of the numerical information (Gamliel & Kreiner, 2019). The present experiment followed the same rule but for the remaining class.

Hypotheses on numeracy were almost supported fully. Firstly, numeracy was the most prominent variable. In general, individuals with higher numeracy performed better in risk evaluation accuracy. High numeracy individuals rated their confidence levels higher and had better risk evaluation accuracy. Participants were equally willing to buy better cars regardless of their numeracy level; high numeracy individuals did not differ in choosing a better car from low numeracy individuals. However, participants low in numeracy were more willing to choose the worse option than high numeracy individuals. This finding is in accordance with the literature (Ghazal et al., 2014; Park & Cho, 2019) that numeracy and financial decisions were associated. The following experiment focused on similar questions with a similar method but with a medical context.

2.3 Experiment 1B

The main scope of the second experiment was to examine individuals' risk perception for medical decisions. Within the same design, the effect of basic knowledge instructions on risk evaluations and willingness to choose the superior option, the relationship between basic instructions and risk evaluations, and willingness ratings were examined. Although numeracy and health decisions were associated concepts, some authors claimed that even though the perceived objective risk is accurate and high, subjective risk perception or belief might impact individuals' final attitudes and behaviors towards better health outcomes (Carey, Herrmann, Hall, Mansfield, & Fakes, 2018; Lau et al., 2022). Therefore, subjective cardiovascular risk and the disease's availability (diagnosis in self or family) were measured in the present experiment. The hypotheses were as follows:H_1: Main effect for visual aids will be observed. Participants who are in the visual aid condition will perform better in risk evaluation accuracy.

H_2: Presence of icon arrays that show part-to-whole relationships will be associated with enhanced confidence levels and willingness to take the medication.

H_3: Willingness to take medication and total risk accuracy will be positively correlated.

H_4: Gist instructions will increase willingness to take the medication more than verbatim and no instruction conditions.

H_5: In the no-instruction condition (when no information is provided that aims at discerning class inclusion), due to numerator's gist pop-out there will be more incorrect responses in risk accuracy; and participants will be underestimating the drug efficacy in risk scenarios (when there is denominator neglect threat).

H_6: An interaction effect is expected to be observed in terms of instruction types and presence of icon arrays that show part-to-whole relationship.

H_7: Interaction effect will be observed between instruction type and numeracy on willingness. Gist instruction, compared to verbatim or no instruction conditions, will yield to more willingness to take medication in low numerate participants than high numerate participants.

H_8: Confidence ratings will be higher in high numerate participants than low numerate participants. This discrepancy will be greater with icon arrays (in visual aid condition).

H_9: Highly numerate participants will perform better in total risk reduction estimates and their willingness will be higher for the better option. Low numerate participants will perform poorer in total risk reduction estimates and their willingness to choose the better option will be lower.

H_{10}: People who have family or self-history about CVD will show increased subjective risk. Subjective risk and willingness to take medication will be positively correlated.

H_{11}: Health literacy and numeracy will be positively correlated.

H_{12}: People who are high in health literacy will show superior willingness to take the medication.

2.3.1 Participants

158 participants took part in the experiment. Exclusion criteria are similar to the previous experiment, with one exception: completing the numeracy scale. Consecutively, 124 (female *N*=82, male *N*=42) were eligible for data analyses. The average age of the participants was 26.6 (*SD* = 9.21, ranging from 18 to 68 years). In general, the sample was highly educated. All participants were at least high school graduates (*N*=58, 46.8%), and the rest were either undergraduates or graduates.

2.3.2 Materials

Similar materials used in both two experiments including the numbers that are used in the scenarios (See Table 2.6), the visuals (See Figure 2.8), and.the instructions (See Table 2.7).

Table 2. 6 Distribution of Numerator and Denominators in Experiment 1B

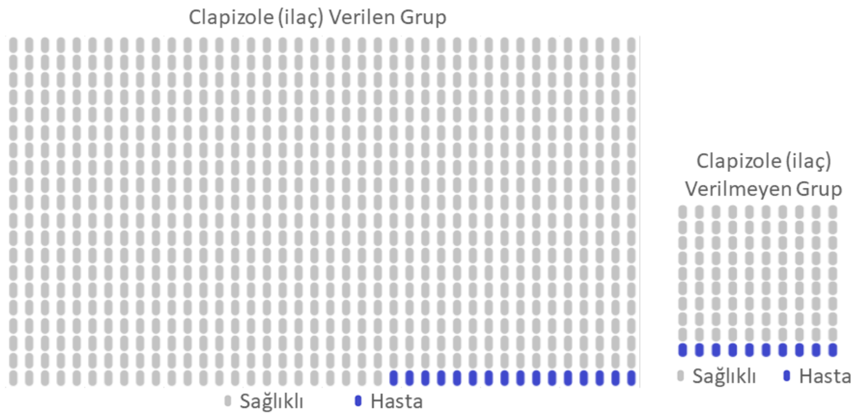
	Medicated		Non-Medicated	
Scenarios and Hypothetical Medicines	Heart Attack	Group Size	Heart Attack	Group Size
Risk (800-100) *Clapizole*	16	800	10	100
Non-risk (100-800) *Tacxolol*	2	100	80	800
Risk (900-150) *Olaxacol*	18	900	15	150
Non-risk (150-900) *Clatolol*	3	150	90	900
Risk (1200-200) *Omdinol*	24	1200	20	200
Non-risk (200-1200) *Topiritol*	4	200	120	1200

Proportions of numerators and denominators are displayed in the table. Scenarios revolved around these numbers and the condition which implied greater amount of faulty cars for the better car numerically is defined as risk evoking scenario.

2.3.2.1 Health literacy

Turkish Health Literacy Scale (Aras & Bayık Temel, 2017) was used in the second experiment (See Appendix A.5). The scale is a measurement of health literacy, which is adapted from (Toçi et al., 2013). The scale is shown as a reliable and valid to use in Turkish adult population with sufficient coefficient (Cronbach's $\alpha = .92$). It includes 25 items in 4 subscales that measures various abilities towards utilization of health information in both cognitive and behavioral level. Subscales are divided as following; (1) ability to access to health information, (2) ability to comprehend information related to health, (3) ability to evaluate (or analyze) and appraise the information related to health, and lastly (4) ability to apply and adhere to the healthy behaviors. Turkish version of this scale includes 25 items which are asked in 5-point Likert scale format (1=I am not able to do it, 2=experiencing great difficulty, 3=experiencing some difficulty, 4=experiencing a little difficulty, 5=no difficulty at all). The maximum score that can be obtained from the scale is 125, and minimum score can be 25. Therefore, high scores in this scale indicated higher subjective health literacy, and low scores indicated lower subjective health literacy.

A. Risk Scenario (800-100)



A. Non-Risk Scenario (100-800)

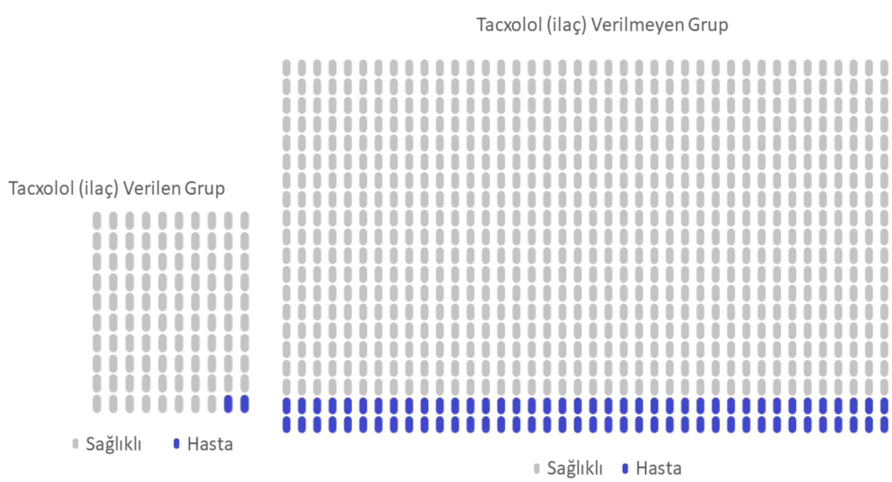


Figure 2. 8 Visuals used in Experiment 1B

Examples of visuals (visual aid condition) used in Experiment 1B for risk (A) and non-risk (B) evoking scenarios. Patients who had heart attack shown in color blue, and healthy individuals are shown in color grey (In non-aided condition, all icons were gray).

Table 2. 7 Instructions used in Experiment 1B

Instruction Type	Scenario Type	Sentence
Gist	For both risk and non-risk	Bu durumda (Clapizole / Tacxolol / Olaxacol / Clatolol / Omdinol / Topiritol) ilacını alan grupta kalp krizi geçirme sıklığı, gerçek ilaç almayan gruptaki kalp krizi geçirme seviyesine göre daha düşüktür.
Verbatim	Risk	Bu durumda Clapizole alan 800 kişiden kalp krizi geçirmeyen kişi sayısı 784, Clapizole almayan 100 kişiden kalp krizi geçirmeyen insan sayısı 90’dır.
		Bu durumda Olaxalol alan 900 kişiden kalp krizi geçirmeyen kişi sayısı 882, Olaxalol almayan 150 kişiden kalp krizi geçirmeyen insan sayısı 135’tir.
		Bu durumda Omdinol alan 1200 kişiden kalp krizi geçirmeyen kişi sayısı 1176, Omdinol almayan 200 kişiden kalp krizi geçirmeyen insan sayısı 180’dir.
	Non-Risk	Bu durumda Tacxolol alan 100 kişiden kalp krizi geçirmeyen kişi sayısı 98, Tacxolol almayan 100 kişiden kalp krizi geçirmeyen insan sayısı 720’dir.
		Bu durumda Clatolol alan 150 kişiden kalp krizi geçirmeyen kişi sayısı 147, Clatolol almayan 900 kişiden kalp krizi geçirmeyen insan sayısı 810’dur.
		Bu durumda Topiritol alan 200 kişiden kalp krizi geçirmeyen kişi sayısı 196, Topiritol almayan 1200 kişiden kalp krizi geçirmeyen insan sayısı 1080’dir.

Table 2.7 shows the instructions that are used in Experiment 1B in each scenario type. All participants saw these instructions under the relevant scenario (risk vs. non-risk)

2.3.2.2 Cardiovascular disease history and subjective risk

Participants were asked to indicate if there is any CVD disease in self or family history. This question was shown in a check box and participants selected yes or no as an answer (to be dummy coded; i.e., 1=Yes, 2=No). In addition to that, subjects were asked to indicate their subjective risk of having a heart condition near future in 10-point single item Likert scale. Lower scores indicated lower risk (See Appendix A.6)

2.3.3 Dependent variables

2.3.2.1 Risk reduction estimate scores

Second experiment took a different approach. Since the scenario included one risk factor with two options of either taking or not taking the medicine, the procedure which is put forward by Schwartz et al. (1997) and used by Okan et al. (2012) was used directly (See Table 2.8). Basically, participants were asked how many individuals would have a heart attack if they take or do not take the medicine in two separate questions. Importantly, since the scenario (or risk evoked by denominator neglect) manipulation was manipulated within subjects so the practice effect was questioned. Therefore, in order to avoid practice effect, the questions were altered at that point. Instead of asking the risk reduction estimate questions all at 1000, the study used 500 in 800-100 and 100-800; 1000 in 900-150 and 150-900 (adhered to Schwartz et al., 1997); 2000 in 1200-200 and 200-1200. The same equation was used in order to calculate risk reduction estimate scores; participants' responses to medicated group are extracted from the responses given to the non-medicated group, then it is divided by 100. The answers were expanded or contracted to the level of the original level (questions were asked in 500 were expanded by 2, and in 2000 were contracted by 2 to the 1000) in order to standardize and center the answers (See Appendix A.7).

Table 2. 8 Calculation of risk evaluations in Experiment 1B

	No Instruction		Gist Instruction		Verbatim Instruction	
RR Questions	Q1	Q2	Q1	Q2	Q1	Q2
Possible Estimations (Raw Risk = E)	E_1= 20	E_2= 100	E_1= 20	E_2= 100	E_1= 20	E_2= 100
Outcome: Correct estimation of the medicine	*=.80.*		*=.80*		*=.80*	
RR Questions	Q1	Q2	Q1	Q2	Q1	Q2
Possible Estimations (Raw Risk = E)	$E_1 \geq 20$	$E_2 <100$	$E_1 \geq 20$	$E_2 <100$	$E_1 \geq 20$	$E_2 <100$
Outcome: Overestimation of the heart attack risk reduced by the medicine	*<.80*		*<.80*		*<.80*	
RR Questions	Q1	Q2	Q1	Q2	Q1	Q2
Possible Estimations (Raw Risk = E)	$E_1 \leq 20$	$E_2 >100$	$E_1 \leq 20$	$E_2 >100$	$E_1 \leq 20$	$E_2 >100$
Outcome: Underestimation of the heart attack risk reduced by the medicine	*>.80*		*>.80*		*>.80*	

The table shows the calculation of risk evaluations across instruction types for both risk and non-risk scenarios. Risk reduction estimate questions were analyzed in two terms in the second experiment. In order to see whether there is an effect of denominator neglect or not (the risk estimates ratio was supposed to be significantly higher (and higher than .80) in risk evoking scenario conditions), Schwartz and colleagues' procedure was followed (1997). Similarly, the answers are coded as correct, underestimation, and overestimation to see the risk accuracy distribution.

2.3.3 Procedure

Participants were randomly assigned into two groups of visually aided condition (icons were showing sub-groups explicitly such as patients who are medicated or non-medicated in color grey, and patients who had heart attack in color blue) and visually non-aided condition (icons were showing the two whole groups in total color grey without showing the sub-groups of patients). In the second part of the experiment, a set of systematic icons depicted the whole group size consisting of individuals who were given the CVD medicine. Patients who had heart attacks shown in the color blue, and healthy individuals are shown in color grey (See Figure 2.8)

There were six scenarios about hypothetical medicine developed by a biomedical company to fight against a risky cardiovascular condition and decrease heart attack risk. Six medicines mentioned in the scenario are named randomly by using a web-based random name generator (https://www.fantasynamegenerators.com). In the scenarios, two groups of people took and did not take medicine (See Appendix A2, A3). The first group was always presented as the group which took medicine, and the second group was always presented as the group which did not take medicine. Similar to the previous experiment, risk reduction across the groups included in the scenarios was kept constant at an 80% rate, which meant that people who took medicine were less likely to have a heart attack by 80%. Participants saw all of six scenarios randomly (See Table 2.3), in counterbalanced order. Instructions' functions were similar, and they were given in a counterbalanced fashion to all participants. Differently, they signaled patients who did not have a heart attack (verbatim instruction) or the advantage of being medicated (gist instruction). Because there were two groups of people who had heart attacks taking or not taking the drug, knowledge instructions highlighted the remaining sub-group of people who did not have the heart attack. It indicated the remaining healthy subjects in both groups who are either medicated or non-medicated in gist and verbatim instruction conditions. There was a difference in the demonstrations of the sentences from the previous experiment. Sentences were written in bold format in order to call participants' attention to the gist and verbatim information (See Table 2.4).

After reading and inspecting the visuals, they were asked to skip the page. And then, participants responded to two risk reduction estimate questions. Risk reduction estimate questions were assigned differently from the first study; instead of asking them all in "X in 1000", the questions were asked in 500 (for 800-100 and 100-800 scenarios) and in 2000 (1200-200 and 200-1200) in order to eliminate a plausible practice effect. In Experiment 1B, instead of open text entry, sliders were utilized for risk reduction estimate questions (open text entry was optional and visible as well on the right side of the scale). Then participants were directed to the next page to rate their confidence levels to their answers (7-point Likert Type Scale, via a slider), followed by the willingness ratings to take medicine on the next page. This loop was repeated for all six scenarios. Participants completed the numeracy and health literacy scales (Aras & Bayık Temel, 2017). On the next page, participants' CVD history was asked, then they were asked to rate their subjective risk or susceptibility towards CVD. They are thanked for their participation and provided the contact information for further questions. The experiment took approximately 25 minutes to complete (see Figure 2.9).

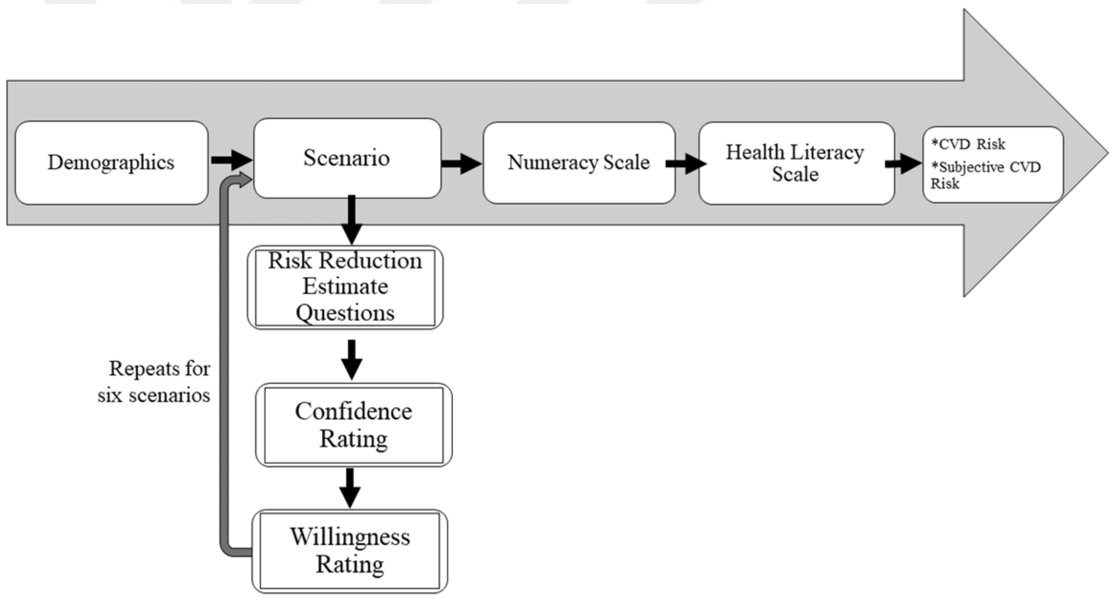


Figure 2. 9 Display Flow of Experiment 1B

Each box refers to a separate page. After filling the demographics form (age, dominant hand, education level, branch of education), an information screen about scenarios was presented and participants are asked to skip the page when they are ready. Scenarios presented in random order without time constraints. The page is skipped then in separate pages participants filled risk reduction estimates (open ended responses), then rated their confidence (on a slider), then rated their willingness (on a slider). This process repeated until all six scenarios are completed. Then participants filled numeracy scale (Lipkus et al., 2001), then health literacy scale. Finally, CVD risk question about self or family history, and subjective CVD risk single item scale was presented.

2.3.4 Results

Firstly, descriptive analysis was conducted. In the study, numeracy scores (*M*=8.46, *SD*=2.35) and health literacy scores (*M*=103, *SD*=15) were generally high across participants who are included in the study (*N*=124). Participants were split into two groups according to their numeracy scores as high (*N*=70, *M*= 10.1, *SD*=0.78) and low (*N*=54, *M*= 6.39, *SD*=2.07); health literacy scores according to their median values as high (*N*=65, *M* = 114, *SD* = 6.41) and low (*N*=57, *M*= 90.8, *SD*=12.6) (median score was 9 for numeracy, and 104 for health literacy).

2.3.4.1. Correlations

Correlation matrix showed that numeracy was positively correlated with health literacy r (122) = .21, p =.02, total correct answers r (119) = .37, p <.001, and overall confidence in answers r (122) = .29, p < .01. Spearman rank correlation showed that numeracy is also correlated with education level r (124) = .18, p < .05. Willingness to take the medication was not correlated with numeracy, r (118) = .16, p = .09 but it was correlated with total correct answers r(115) = .32, p <.001.

2.3.4.2 Risk reduction estimates

2x3x2x2 mixed factorial ANOVA was conducted for two within subject factors; scenario types (risk vs. non-risk) and instruction types (no vs. gist vs. verbatim) (scenario types, instruction types, and visual aid), two between subject factors (numeracy and visual) on three dependent variables (risk reduction estimates, willingness, and confidence). Risk reduction estimates were calculated via the formula which is adapted from Schwartz et al. (1997) and Okan et al. (2011). Analyses did not indicate main effect for using visual aids on risk reduction estimates F (1, 117) = 1.72, p = .192, $\eta^2 p$ =.014. Similarly, numeracy and instruction types did not show significant effect on risk reduction estimates; respectively F $(1, 117) = 1.57$, $p = .212$, $\eta^2p = .013$, $F (1.85, 216.61) = 1.62$, $p = .203$, $\eta^2p = .014$. However, scenario type showed a main effect $F (1, 117) = 9.63$, $p = .002$, $\eta^2p = .076$. The interaction effect between numeracy and visual aid was not significant (See Figure 2.10).

2.3.4.3 Risk reduction accuracy

Similar to the previous study, participants' responses to risk reduction estimates were coded as accurate, and not accurate (underestimation and overestimation of the drug efficacy) (See Figure 3.2). Results showed that only numeracy had the main effect on risk reduction accuracy $F(1, 117) = 9.25$, $p = .003$, $\eta^2p = .073$, favoring towards individuals with higher numeracy (See Figure 2.11). There was no significant main effect of scenario type, instruction type, visuals on risk reduction accuracy. None of the interactions were significant.

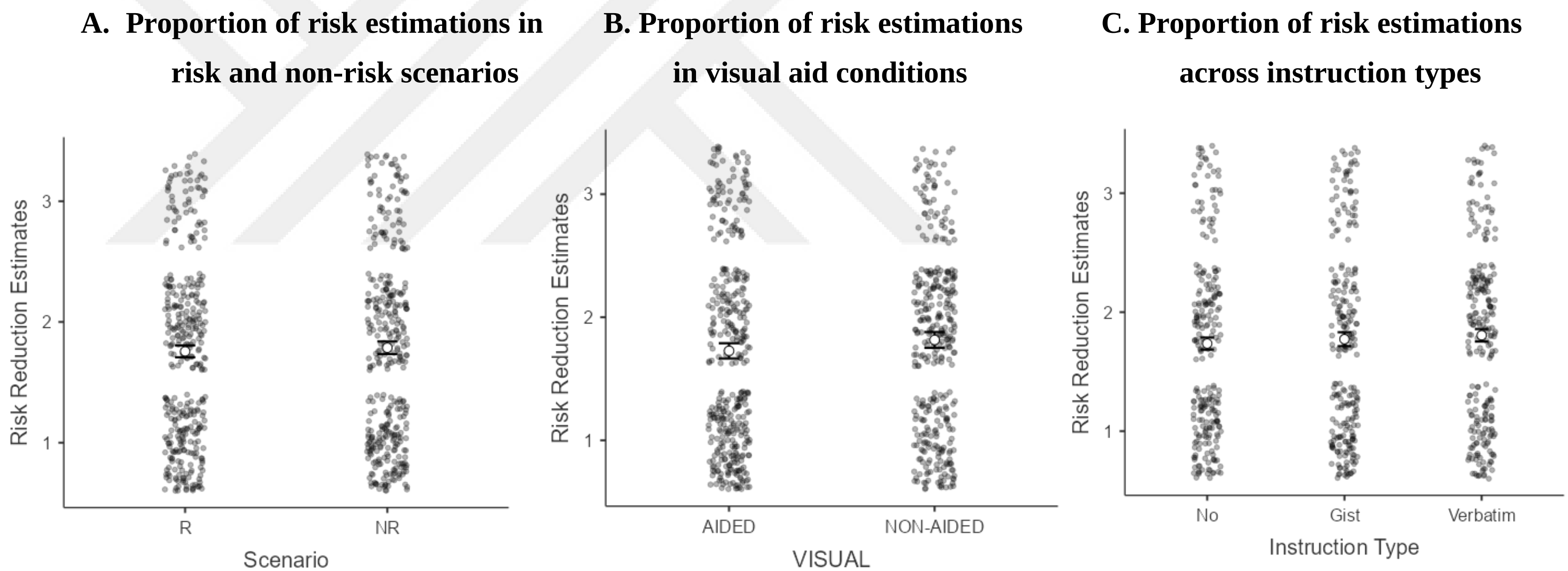
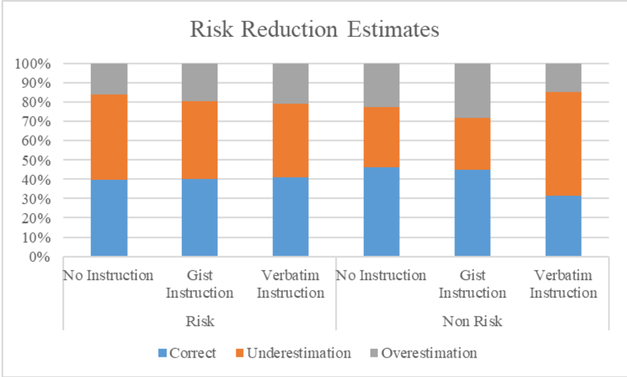


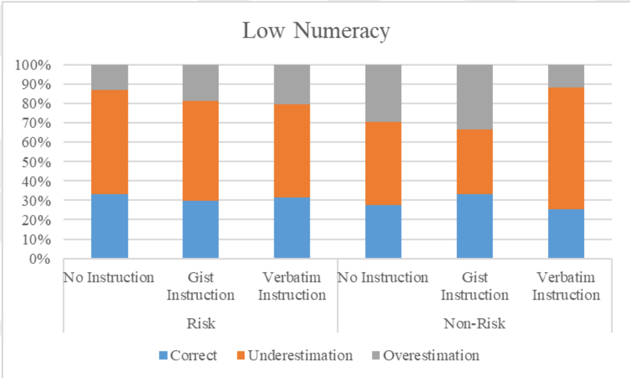
Figure 2.10 Distributions of Correct and Incorrect Responses in Experiment 1B

First graph (A) shows participants' correct, and incorrect risk estimates (underestimation and overestimation of the risk reduced by the medicine) when the scenario indicated risk (R) or not (NR). Second graph (B) shows the same risk estimates in visually aided or non-aided conditions. Third graph (C) shows the risk estimates when the scenarios are given in three different instruction formats (no, gist, verbatim). Numbers on Y-axis (i.e., 1, 2, 3) represented correct (1), overestimation of the risk reduced by the medicine (2), and underestimation of the risk reduced by the medicine (3). Dots represented observed scores of the participants. Error bars represented one standard error. Accumulation of dots in 1 shows the frequency of the responses that are correct, in 2 that are incorrect and overestimated, in 3 that are incorrect and underestimated.

A. Risk Estimations Across Instruction Types



B. Risk Estimations with Low Numeracy



C. Risk Estimations with High Numeracy

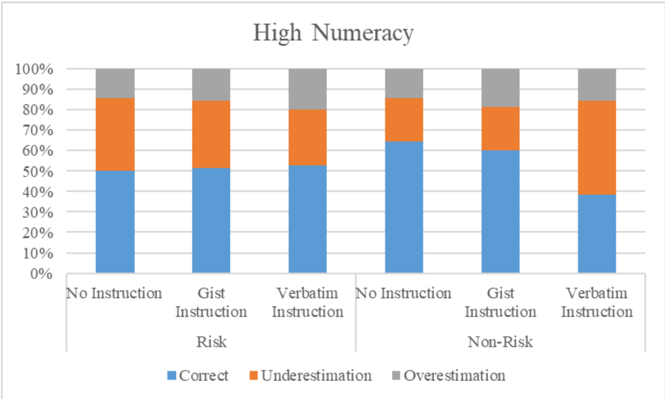


Figure 2.11 Percentages of Correct and Incorrect Answers in 1B

The graphs show the percentages of participants whose answers are accurate, or inaccurate (status) as the function of scenario and instruction type manipulations for participants with both low (B), and high (C) numeracy.

2.3.4.4. Confidence ratings

Numeracy has the main effect on confidence levels F (1, 118) = 15.01, $p < .001$, η^2p =.113. Individuals with higher numeracy (M = 3.87, SE = .28) rated their confidence levels higher in general. None of the other factors achieved statistical significance. Visual aids, instruction types, and scenario type had no main effect on confidence levels. Although the interaction was not significant, for exploratory analyses, post-hoc comparison was examined for the interaction between visual aid and numeracy, F (1, 118) = 3.10, p = .081, η^2p =.026. When there is no visual aid, confidence ratings were significantly different between individuals with low (M = 3.87, SE = .28, 95% CI [3.33, 4.42]) and high numeracy (M= 5.37, SE = .26, 95% CI [4.68, 5.88]), p < .001 (See Figure 2.11).

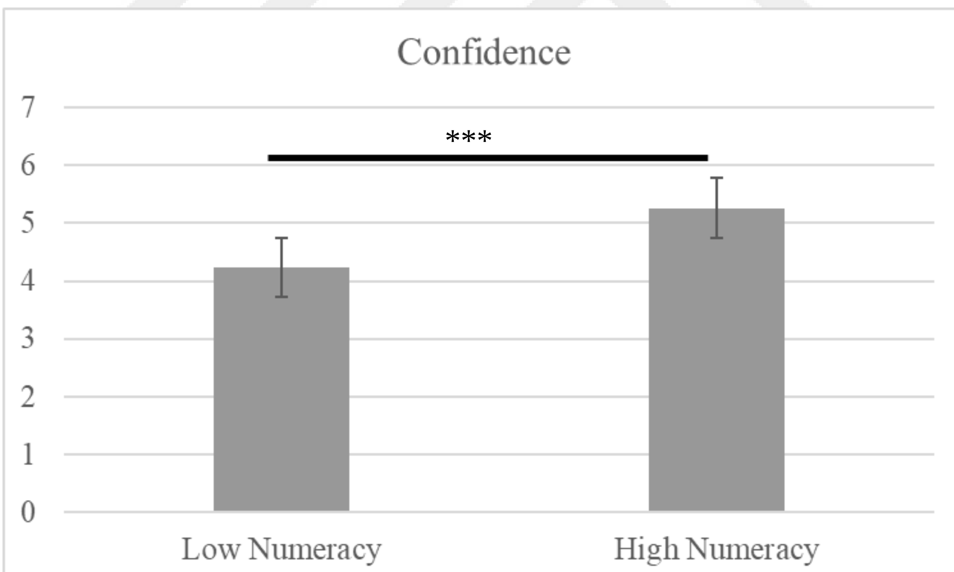


Figure 2.12 Confidence Ratings of Participants in Experiment 1B

The graph shows the confidence in responses to risk reduction estimate questions between low and high numeracy participants. Error bars stand for one standard error (***p <.001)

2.3.4.5. Willingness ratings

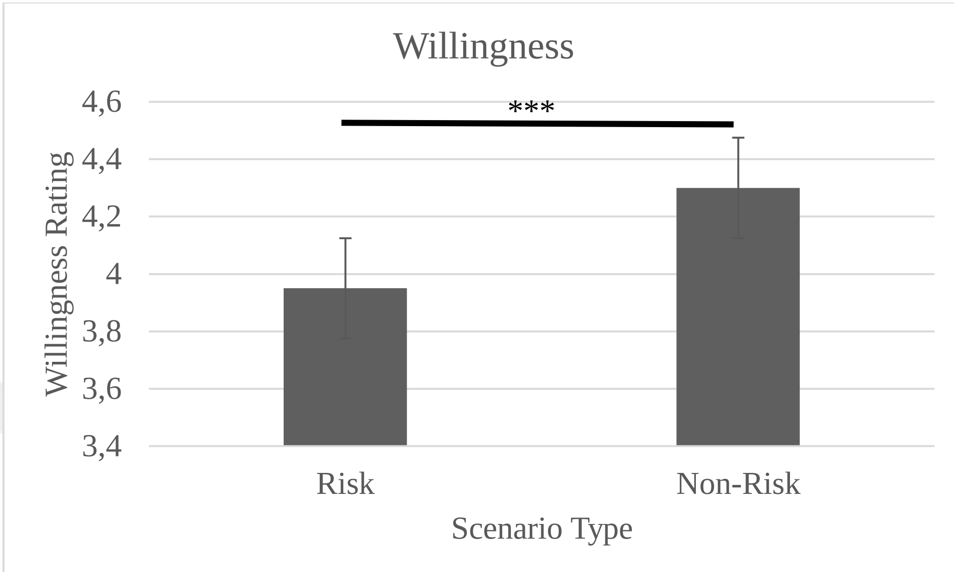
Mixed factorial ANOVA was conducted with two within subject factors (instruction types and scenario types) and three between subject factors (visual aid, numeracy, and health literacy) on willingness. Results showed that health literacy did not have the main effect on willingness to use the medication F (1, 111) = .12, p = .70, $\eta^2 p$ =.001. Numeracy showed a main effect on willingness F (1, 111) = .12, p = .02, $\eta^2 p$ =.052 (See Figure 1.13B).

Scenario type has the main effect on willingness F (1, 111) = 11.31, p = .001, $\eta^2 p$ =.095. Risk evoking scenarios (*M* = 3.95, *SE* = .16, 95% CI [3.64, 4.26]) yielded participants to rate their willingness lower compared to non-risk evoking scenarios (*M* = 4.30, *SE* = .16, 95% CI [3.98, 4.62]) (See Figure 2.13A). The two-way interaction between scenario type and numeracy was also significant F (1, 111) = 10.18, p = .003, $\eta^2 p$ =.079. When the scenarios are risk evoking, participants with lower numeracy rated their willingness lower than when the scenarios are not risk evoking. Three-way interaction between scenario type, visual, and numeracy was not statistically significant F (1, 111) = 10.18, p = .052, $\eta^2 p$ =.034.

2.3.4.6 CVD risk and subjective risk

Independent samples t-test was revealed that having a CVD history was associated with subjective risk perception $t(118) = 6.50$, $p < .001$. Participants with higher availability for the disease, who have self or family history with CVD, rated their subjective risk higher (*M* = 5.24, *SD* = 2.44) than participants did not have CVD history (*M* = 2.58, *SD* = 1.89).

A. **Willingness Ratings Across Scenario Conditions**



B. **Willingness Ratings (Scenario*Numeracy)**

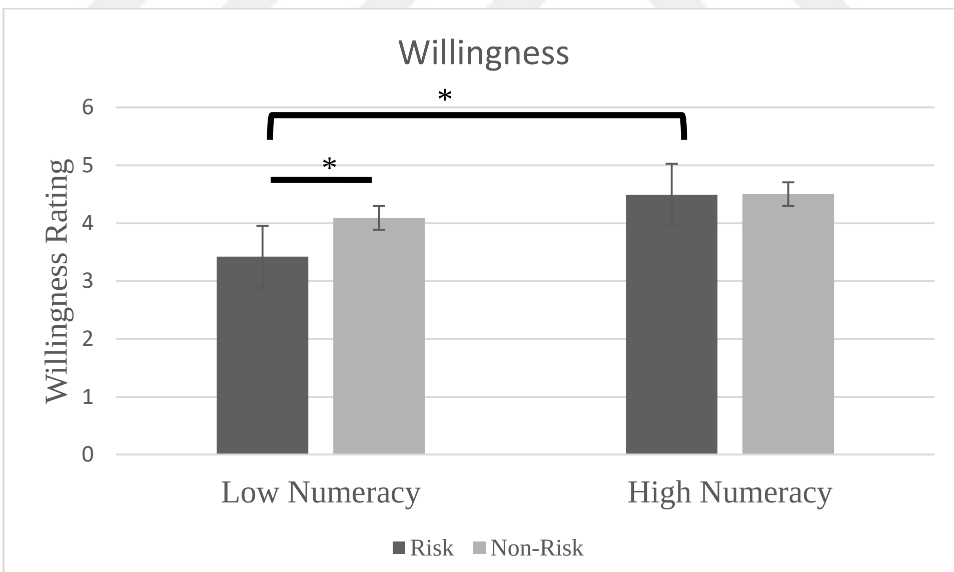


Figure 2.13 Willingness Ratings of Participants in Experiment 1B

The graphs show the willingness to use the medication across scenario types (A) and function of scenario and numeracy interaction low and high numeracy participants (B). Error bars stand for one standard error. (*p<.05, **p <.01***p<.001)

3.1.5 Discussion

The aim was to inspect the effects of numeracy, visual aids, knowledge types, and instruction types on risk evaluations, confidence levels, and willingness in the medical context. Denominator neglect was observed in estimates; risk reduction estimates were affected when the risk was presented in risk-evoking scenarios. Participants underestimated the drug's risk reduction as the function of scenarios that are risk-evoking or not. Similarly, when the risk is presented in the non-risk evoking scenario, participants overestimated the risk reduction of the drug. More importantly and contrary to the first hypothesis, visual aids did not help to reduce the denominator neglect. The only mitigating factor for denominator neglect was found to be numeracy. Numeracy yielded better risk estimations and, therefore, better risk perception accuracy. While this finding overlaps with most of the previous field studies, it contradicts nearly the same amount by not showing the graphical displays' improvement on risk evaluations. Moreover, this finding contradicts Experiment 1A, in which the results did not show any main effect for scenario types on risk evaluations.

Instruction types had no main effect on risk estimates even though the priming of the sentences was altered (discussed in the previous study's discussion part). Furthermore, none of the interactions yielded statistically significant results. Gist instruction was no different from the remaining conditions and the no-instruction condition. Therefore, this finding claimed that knowledge evoking should be considered in terms of the method rather than participants' motivational or attentional factors to focus on instructions.

The hypothesis about the relationship between confidence ratings and numeracy was partially supported. Indeed, participants with high numeracy were more confident about their risk reduction estimate responses. Contrarily to the consecutive hypothesis, when there was no visual aid, this discrepancy increased between individuals with low and high numeracy. Although low numeracy is associated with less risk reduction accuracy independent from visual aids condition, it is also associated with better confidence when the risk is presented with visual aids. This gives an interesting insight because individuals with high and low numeracy can equally feel confident about their responses when there is a visual aid, while numeracy still affects the risk estimate accuracy. When the numeracy is low, the accuracy of risk estimations drops, although the confidence is high. Low numeracy individuals feel more confident about their answers when visual aids are presented; however, that confidence was not supported with accuracy. When the numeracy is low, the accuracy of risk estimations drops, although the confidence is high. This is also an interesting challenge for subjective numeracy arguments, which claim the need for the coexistence of objective and subjective numeracies (Peters & Bjalkebring, 2015).

Health literacy was correlated with total risk accuracy, numeracy, and education level, except for the willingness to take the medication. In addition, health literacy had no significant effect on willingness ratings. These two findings on willingness were in contradiction with the hypotheses. While this overlaps with Golbeck and colleagues' (2005) view of health numeracy (numerical understanding of health information rather than reflective-behavioral axis), it is also a contradictory finding for health literacy as the adherence to and motivation for making health decisions (Aras & Bayık Temel, 2017). Besides the discussions over the operationalization of health literacy which focuses on whether it is an analytical or more of a practical construct, or an individual or more of a collective construct (See for a brief review, Berkman et al., 2010), present results showed that willingness to use the medication (which is a practical and individual construct in the present experiment) was not associated with numerical understanding of the risks (while it is associated with numerical abilities or objective numeracy in general). Results supported that people with a family history or self-history of CVD demonstrated increased subjective risk (exploratory analysis within the same design, CVD history as the between-subject variable, did not yield a significant effect of dependent variables and controlling for subjective risk).

3. GENERAL DISCUSSION

The importance of accurate risk communication in various fields is becoming more evident in recent years. Especially while the world was facing a sudden pandemic crisis and dealing with a massive flow of complex information, the individual and society's risk understanding was facing great difficulty. (Lachlan, Hutter, Gilbert, & Spence, 2021). Moreover, the accurate information given by public authorities were questioned because of the way they were given (Jie, 2022), and visualization techniques were used (Comba, 2020) without knowing the psychological factors that might affect the risk perceptions and yield misleading (i.e., using visuals that are complex to understand, or using tables that gives daily cases fused with denominator neglect). Therefore, the essentiality of risk communication was under the scope of the decision and risk researchers (Lachlan, Hutter, Gilbert, & Spence, 2021; Lau et al., 2022; Reyna et al., 2021; Wolfe, 2021). The present thesis examined the effects of visual aids and knowledge evoking types within complex risk scenarios in non-medical and medical contexts in two experiments. The denominator neglect effect, which can be detrimental to adapting superior choices due to the inaccurate (enhanced) risk perception in risk-evoking situations, was not observed in non-medical scenarios. However, it was observed in the medical scenarios. However, unlike the previous literature, denominator neglect is not reduced via icon arrays as visuals (Garcia-Retamero, Okan, & Cokely, 2012; Okan et al., 2011). Results showed that visual aids as icon arrays did not help individuals in terms of accurate risk estimations, that is, whether the scenario indicated risk or not. Instruction types, which are formed via the fuzzy trace theory's approach, did not significantly impact the accuracy of risk evaluations.

The denominator neglect effect is replicated in the second experiment; in risk scenarios, medication's efficacy was underestimated (the same pattern is observed in the first experiment; however it did not reach statistical significance). In other words, denominator neglect was more salient in medical scenarios. However, this finding should be taken with a grain of salt, because there are no supportive findings or no hypotheses that are considered in this paper to explain this difference between two experiments. Moreover, there are studies that show the similar patterns in medical (Okan et al., 2011) and non-medical scenarios (Chua, Yates, & Shah, 2006).

This difference might have arisen due to the alterations in data collection processes between the two experiments. Firstly, changing risk estimate questions were first to think because the indicator of the denominator neglect was the risk reduction accuracy which was derived from risk reduction estimate scores in both experiments. Altering them might sound problematic initially because it creates widened or shrunk base rates for participants to respond. However, this cannot explain the difference between two experiments, because of the counterbalanced design of scenario types (questions that have wide gap can be asked under the other within subject variables, any dyad of numbers and risk reduction estimate questions can be encountered in any instruction type and scenario type). Therefore, systemic error if there is, is distributed across conditions. Moreover, the absence of the denominator neglect effect in Experiment 1A might be due to the practice effects since participants might have figured out the correct answers. Therefore, reducing the practice effect by changing risk reduction estimate questions can be attained for denominator neglect in 1B (participants needed to alter their risk reduction estimate answers since the questions were not asked in the same number in Experiment 1B).

The difference between text boxes and slider scales might have also yielded these results (Thomas & Kyung, 2019). Utilizing sliders (which shows proportional area between the slider's indicating point and the remaining point from zero to the attributed risk estimate response instead of text entries which does not indicate any visual cue); and the numbers that are chosen for the risk estimate questions that can create a visually asymmetric gap between the scenarios. Although the length of the sliders was visually the same and was cut into the equally proportionate scale indicators in 1B, participants should have dragged the range selector into responding accurately to a similar near left-end point (heart attack rate for non-medicated 10-20-40 or heart attack rate for non-medicated 50-100-200 in 500-1000-2000 scales, respectively), which leaves a great deal of space in different risk reduction estimate questions. Although there was an open text box next to the slider (optional if the slider's range selector does not work properly), there is no data for participants who used or did not use the open text box. On the other hand, as claimed in the discussion part of the first experiment, comparative options might be understood differently than the bound options included in the second study (1B: taking or not taking medicine vs. 1:A buying a better or worse car). In other words, even though the scenarios indicated the same numbers with the same risk ratios, the comparative nature of the first experiment and adjustments of the answers regarding the other option might be playing a role.

To sum up, some studies suggest that visual aids can be detrimental for short recall of the information after minutes of presentation (Ruiz et al., 2013). However, that was not the case in this study since participants saw the risk estimate questions right after the demonstrations in both experiments. In addition, in exploratory analyses, when CVD history is put into the analyses, the results did not change; therefore, these findings cannot be explained with risk as an affective factor (Wilhelms, Fraenkel, & Reyna, 2018). Numeracy was the strongest predictor for risk evaluations, confidence, and willingness toward better choices or better health decisions in accordance with previous findings (Helm, Hans, Reyna, & Reed, 2020; Låg, Bauger, Lindberg, & Friborg, 2014; Peters, 2008; Reyna & Brainerd, 2007; Sullivan et al., 2021). The sample was generally high in numeracy. Considering that high numerate individuals were also good at the computation of the risk indicators whether the risk is given with the numbers only or with graphical demonstrations (Fraenkel et al., 2018; Oudhoff & Timmermans, 2015). However, the results were symmetric when two experiments were compared, and numeracy was similarly affecting risk accuracy. Therefore, the detrimental effects of visual aids, affective factors, and numerical competency were ruled out when explaining the difference between the two experiments.

In short, the difference in the results might be examined for three possible reasons, the changes in experimental set-up (text entry was removed, sliders replaced them), which might be affected from the visual cue created by sliders; changes in risk reduction estimate questions that reduces the practice effect in within-subject manipulation of scenario type (Koehler, 1996) (instead of using the same denominator in the Schwartz et al. (1997) study, changing risk reduction estimate questions to 500,1000, and 2000), and the difference between two tasks' nature.

3.1 Limitations and Strengths

The first limitation of the present study was the stimuli that were chosen. Firstly, the material that is used in the study is theoretically accredited. For instance, instruction types were formed by the gist communication approach (Reyna, 2012), as well as scenario types formed by leaning on broad literature on denominator neglect and its main findings, which was the overestimation of the risks when the numerator is greater. This has great advantages, such as having a theoretical ground for the outcome and data interpretation. On the other hand, the lack of pilot study in terms of scenario type and instruction types is one of the points that need improvement in the present study. Secondly, the sample consisted of well-educated, and they were generally high in numeracy. Since the results consistently indicated the main effect of numeracy, and the connection between literacy and numeracy was known and discussed, this should bear a more cautious interpretation of the data. The second limitation was the sample. The present study's scores obtained from Objective Numeracy Scale (Lipkus et al., 2001) were generally high across participants. In addition to the great numeric ability in the sample, participants were also highly educated (all included participants were at least high school graduates and were continuing their studies at the bachelor's or associate degree level).

Strengths can be summarized as the new approach to knowledge evoking and graph interaction. Previous studies were trying to explain the knowledge that is inferred from graphical or numerical risk representations rather than the mechanism and nature of the relationship. This study, although it failed to explain the association, is one of the first ones to examine knowledge evoking instead of knowledge extraction.

In addition to that, the study examined risk evaluations in two different contexts (medical and non-medical) using the same methodology and nearly the same materials. Therefore, results can be interpreted comprehensively regarding financial and medical concepts.

However, that does not mean that disjointing nested classes via knowledge evoking is not a method that cannot be utilized. On the contrary, it can be achieved via analytic thinking training about classes and subclasses (Brainerd & Reyna, 1990; Wolfe & Reyna, 2010).

3.2 Future Implications

As discussed in the limitation section, since the single instruction sentences did not have a significant effect in this study, knowledge evoking might be more complicated than single sentences for risk judgments, and it needs improvement in terms of taking different approaches rather than verbal. Analytic thinking training can be one of these approaches in terms of the gist or verbatim training. Although icon arrays are shown to aid gist extraction in literature, the mechanism between graphical demonstration and knowledge inference remained unclarified. Secondly, the demonstration of instructions might be altered, and the absence of the effects of instructions might be revisited by manipulating the instruction-scenario congruency. For example, when there is an incongruence in information between scenario and instruction, the presence of an effect would result in less accuracy and confidence, and the true absence of instruction effect would not yield a significant difference. Moreover, even if the participants would be similarly accurate or confident, variance in the reaction times can be used to detect the impact of instructions between congruent and incongruent information.

REFERENCES

Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. B. (2006). Design Features of Graphs in Health Risk Communication: A Systematic Review. *Journal of the American Medical Informatics Association*, *13*(6), 608–618. https://doi.org/10.1197/jamia.M2115

Aras, Z., & Bayık Temel, A. (2017). Sağlık Okuryazarlığı Ölçeğinin Türkçe Formunun Geçerlilik Ve Güvenirliğinin Değerlendirilmesi (Evaluation of the Validity and Reliability of the Turkish Form of the Health Literacy Scale). *Florence Nightingale Hemşirelik Dergisi*, *25*(2), 85.

Ashby, N. J. S., Rakow, T., & Yechiam, E. (2017). 'Tis better to choose and lose than to never choose at all. *Judgment and Decision Making*, *12*(6), 553–562.

Attema, A. E., L'Haridon, O., Raude, J., & Seror, V. (2021). Beliefs and Risk Perceptions About COVID-19: Evidence From Two Successive French Representative Surveys During Lockdown. *Frontiers in Psychology*, *12*(February), 1–16. https://doi.org/10.3389/fpsyg.2021.619145

Bailey, R., Willner, P., & Dymond, S. (2011). A visual aid to decision-making for people with intellectual disabilities. *Research in Developmental Disabilities*, *32*(1), 37–46. https://doi.org/10.1016/j.ridd.2010.08.008

Bancilhon, M., & Ottley, A. (2020). Did you get the gist of it? understanding how visualization impacts decision-making. *ArXiv*, 12–14.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*(3), 241–297. https://doi.org/10.1017/S0140525X07001653

Berkman, N. D., Davis, T. C., & McCormack, L. (2010). Health literacy: What is it? *Journal of Health Communication*, *15*(SUPPL. 2), 9–19. https://doi.org/10.1080/10810730.2010.499985

Blalock, S. J., & Reyna, V. F. (2016). Using fuzzy-trace theory to understand and improve health judgments, decisions, and behaviors: A literature review. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *35*(8), 781–792. https://doi.org/10.1037/hea0000384

Bonner, C., Trevena, L. J., Gaissmaier, W., Han, P. K. J., Okan, Y., Ozanne, E., … Zikmund-fisher, B. J. (2021). *Current Best Practice for Presenting Probabilities in Patient Decision Aids : Fundamental Principles*. 1–13. https://doi.org/10.1177/0272989X21996328

Brainerd, C. J., & Reyna, V. F. (1990). Inclusion illusions: Fuzzy-trace theory and perceptual salience effects in cognitive development. *Developmental Review*, *10*(4), 365–403. https://doi.org/10.1016/0273-2297(90)90020-5

Broniatowski, D. A., & Reyna, V. F. (2020). To illuminate and motivate: a fuzzy-trace model of the spread of information online. In *Computational and Mathematical Organization*

Theory (Vol. 26). https://doi.org/10.1007/s10588-019-09297-2

Carey, M., Herrmann, A., Hall, A., Mansfield, E., & Fakes, K. (2018). Exploring health literacy and preferences for risk communication among medical oncology patients. *PLoS ONE*, *13*(9), 1–12. https://doi.org/10.1371/journal.pone.0203988

Cho, Y. H., You, M., & Choi, H. (2018). Gist-based design of graphics to reduce caffeine consumption among adolescents. *Health Education Journal*, *77*(7), 778–790. https://doi.org/10.1177/0017896918765024

Chua, H. F., Yates, J. F., & Shah, P. (2006). Risk avoidance: Graphs versus numbers. *Memory and Cognition*, *34*(2), 399–410. https://doi.org/10.3758/BF03193417

Comba, J. L. D. (2020). Data Visualization for the Understanding of COVID-19. *Computing in Science and Engineering*, *22*(6), 81–86. https://doi.org/10.1109/MCSE.2020.3019834

Copurlar, C., & Kartal, M. (2016). What is Health Literacy? How to measure it? Why is it important? *Turkish Journal of Family Medicine & Primary Care*, *10*(1), 40. https://doi.org/10.5455/tjfmpc.193796

Dodd, R. H., Cvejic, E., Bonner, C., Pickles, K., McCaffery, K. J., Ayre, J., … Nickel, B. (2021). Willingness to vaccinate against COVID-19 in Australia. *The Lancet Infectious Diseases*, *21*(3), 318–319. https://doi.org/10.1016/S1473-3099(20)30559-4

Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2011). Helping patients decide: Ten steps to better risk communication. *Journal of the National Cancer Institute*, *103*(19), 1436–1443. https://doi.org/10.1093/jnci/djr318

Feldman-Stewart, D., Brundage, M. D., & Zotov, V. (2007). Further insight into the perception of quantitative information: Judgments of gist in treatment decisions. *Medical Decision Making*, *27*(1), 34–43. https://doi.org/10.1177/0272989X06297101

Felicity Goodyear-Smith, Bruce Arroll, Lydia Chan, Rod Jackson, Sue Wells, Timothy Kenealy. (2008). Patients prefer pictures to numbers to express cardiovascular benefit from treatment (Annals of Family Medicine (2008)). *Annals of Family Medicine*, *6*(4), 372. https://doi.org/10.1370/afm.881

Fraenkel, L., Reyna, V., Cozmuta, R., Cornell, D., Nolte, J., & Wilhelms, E. (2018). Do visual aids influenced patients' risk perceptions for rare and very rare risks? *Patient Education and Counseling*, *101*(11), 1900–1905. https://doi.org/10.1016/j.pec.2018.06.007

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. https://doi.org/10.1037/a0024850

Gamliel, E., & Kreiner, H. (2019). Applying Fuzzy-Trace Theory to Attribute-Framing Bias: Gist and Verbatim Representations of Quantitative Information. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(3), 497–506.

https://doi.org/10.1037/xlm0000741

Garcia-Retamero, R., & Dhami, M. K. (2011). Pictures speak louder than numbers: On communicating medical risks to immigrants with limited non-native language proficiency. *Health Expectations*, *14*(SUPPL. 1), 46–57. https://doi.org/10.1111/j.1369-7625.2011.00670.x

Garcia-Retamero, R., Galesic, M., & Gigerenzer, G. (2010). Do icon arrays help reduce denominator neglect? *Medical Decision Making*, *30*(6), 672–684. https://doi.org/10.1177/0272989X10369000

Garcia-Retamero, R., Okan, Y., & Cokely, E. T. (2012). Using visual aids to improve communication of risks about health: A review. *The Scientific World Journal*, *2012*. https://doi.org/10.1100/2012/562637

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.

Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine*, *29*(4), 375–376. https://doi.org/10.1016/j.amepre.2005.06.012

Hawley, S. T., Zikmund-Fisher, B., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling*, *73*(3), 448–455. https://doi.org/10.1016/j.pec.2008.07.023

Helm, R. K., Hans, V. P., Reyna, V. F., & Reed, K. (2020). Numeracy in the jury box: Numerical ability, meaningful anchors, and damage award decision making. *Applied Cognitive Psychology*, *34*(2), 434–448. https://doi.org/10.1002/acp.3629

Hess R., Vischers V. H. M., S. M. (2011). Risk communication with pictographs: The role of numeracy and graph processing. *Judgment and Decision Making*, *6*(3), 263–274.

Hux, J. E., & Naylor, C. D. (1995). Communicating the Benefits of Chronic Preventive Therapy: Does the Format of Efficacy Data Determine Patients' Acceptance of Treatment? *Medical Decision Making*, *15*(2), 152–157. https://doi.org/10.1177/0272989X9501500208

Jie, Y. (2022). Frequency or total number? A comparison of different presentation formats on risk perception during COVID-19. *Judgment and Decision Making*, *17*(1), 215–236.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, *19*(1), 1–53. https://doi.org/10.1017/s0140525x00041157

Kurz-Milcke, E., Gigerenzer, G., & Martignon, L. (2008). Transparency in risk communication: Graphical and analog tools. *Annals of the New York Academy of*

Sciences, *1128*, 18–28. https://doi.org/10.1196/annals.1399.004

Lachlan, K. A., Hutter, E., Gilbert, C., & Spence, P. R. (2021). From what I've heard, this is bad: An examination of Americans' source preferences and information seeking during the COVID-19 pandemic. *Progress in Disaster Science*, *9*, 100145. https://doi.org/10.1016/j.pdisas.2021.100145

Låg, T., Bauger, L., Lindberg, M., & Friborg, O. (2014). The Role of Numeracy and Intelligence in Health-Risk Estimation and Medical Data Interpretation. *Journal of Behavioral Decision Making*, *27*(2), 95–108. https://doi.org/10.1002/bdm.1788

Lau, N. T. T., Wilkey, E. D., Soltanlou, M., Lagacé Cusiac, R., Peters, L., Tremblay, P., … Ansari, D. (2022). Numeracy and COVID-19: examining interrelationships between numeracy, health numeracy and behaviour. *Royal Society Open Science*, *9*(3). https://doi.org/10.1098/rsos.201303

Levy, H., Ubel, P. A., Dillard, A. J., Weir, D. R., & Fagerlin, A. (2014). Health numeracy: The importance of domain in assessing numeracy. *Medical Decision Making*, *34*(1), 107–115. https://doi.org/10.1177/0272989X13493144

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making*, *27*(5), 696–713. https://doi.org/10.1177/0272989X07307271

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. https://doi.org/10.1177/0272989X0102100105

Navarro, J. D., & Foxcroft, D. R. (2019). Learning statistics with jamovi: A tutorial for psychology students and other beginners (Version 0.70). *Journal of Chemical Information and Modeling*. Retrieved from https://www.learnstatswithjamovi.com/

Nayak, J. G., Hartzler, A. L., Macleod, L. C., Izard, J. P., Dalkin, B. M., & Gore, J. L. (2016). Relevance of graph literacy in the development of patient-centered communication tools. *Patient Education and Counseling*, *99*(3), 448–454. https://doi.org/10.1016/j.pec.2015.09.009

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2011). Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension. *Journal of Behavioral Decision Making*, *25*(4), 390–401. https://doi.org/10.1002/bdm.751

Oliver, A. (2003). A quantitative and qualitative test of the Allais paradox using health outcomes. *Journal of Economic Psychology*, *24*(1), 35–48. https://doi.org/10.1016/S0167-4870(02)00153-8

Ooms, J. A., Jansen, C. J. M., & Hoeks, J. C. J. (2020). The story against smoking: An exploratory study into the processing and perceived effectiveness of narrative visual smoking warnings. *Health Education Journal*, *79*(2), 166–179. https://doi.org/10.1177/0017896919867436

Oudhoff, J. P., & Timmermans, D. R. M. (2015). The effect of different graphical and numerical likelihood formats on perception of likelihood and choice. *Medical Decision Making*, *35*(4), 487–500. https://doi.org/10.1177/0272989X15576487

Park, I., & Cho, S. (2019). The influence of number line estimation precision and numeracy on risky financial decision making. *International Journal of Psychology*, *54*(4), 530–538. https://doi.org/10.1002/ijop.12475

Peters, E. (2008). Numeracy and the perception and communication of risk. *Annals of the New York Academy of Sciences*, *1128*, 1–7. https://doi.org/10.1196/annals.1399.001

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology*, *108*(5), 802–822. https://doi.org/10.1037/pspp0000019

Peters, E., Tompkins, M. K., Knoll, M. A. Z., Ardoin, S. P., Shoots-Reinhard, B., & Meara, A. S. (2019). Despite high objective numeracy, lower numeric confidence relates to worse financial and medical outcomes. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(39), 19386–19391. https://doi.org/10.1073/pnas.1903126116

Quillian, L., & Pager, D. (2010). Estimating risk: Stereotype amplification and the perceived risk of criminal victimization. *Social Psychology Quarterly*, *73*(1), 79–104. https://doi.org/10.1177/0190272509360763

Reyna, V. (2018). When irrational biases are smart: A fuzzy-trace theory of complex decision making. *Journal of Intelligence*, *6*(2), 1–16. https://doi.org/10.3390/jintelligence6020029

Reyna, V. F. (2012). Risk perception and communication in vaccination decisions: A fuzzy-trace theory approach. *Vaccine*, *30*(25), 3790–3797. https://doi.org/10.1016/j.vaccine.2011.11.070

Reyna, V. F., & Brainerd, C. J. (1991). Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making*, *4*(4), 249–262. https://doi.org/10.1002/bdm.3960040403

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

Reyna, Valerie F. (1991). Class inclusion, the conjunction fallacy, and other cognitive illusions. *Developmental Review*, *11*(4), 317–336. https://doi.org/10.1016/0273-2297(91)90017-I

Reyna, Valerie F. (2020). A scientific theory of gist communication and misinformation resistance, with implications for health, education, and policy. *Proceedings of the National Academy of Sciences*, 201912441. https://doi.org/10.1073/pnas.1912441117

Reyna, Valerie F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences*, *17*(2), 147–159.

https://doi.org/10.1016/j.lindif.2007.03.010

Reyna, Valerie F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89–107. https://doi.org/10.1016/j.lindif.2007.03.011

Reyna, Valerie F., & Brainerd, C. J. (2014). Fuzzy universality of probability judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(48), 16984–16985. https://doi.org/10.1073/pnas.1420118111

Reyna, Valerie F., Broniatowski, D. A., & Edelson, S. M. (2021). Viruses, Vaccines, and COVID-19: Explaining and Improving Risky Decision-making. *Journal of Applied Research in Memory and Cognition*, *10*(4), 491–509. https://doi.org/10.1016/j.jarmac.2021.08.004

Reyna, Valerie F., & Brust-Renck, P. G. (2020). How representations of number and numeracy predict decision paradoxes: A fuzzy-trace theory approach. *Journal of Behavioral Decision Making*, *33*(5), 606–628. https://doi.org/10.1002/bdm.2179

Reyna, Valerie F., Chick, C. F., Corbin, J. C., & Hsia, A. N. (2014). Developmental Reversals in Risky Decision Making: Intelligence Agents Show Larger Decision Biases Than College Students. *Psychological Science*, *25*(1), 76–84. https://doi.org/10.1177/0956797613497022

Reyna, Valerie F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, *12*(3), 179–195. https://doi.org/10.1037/1076-898X.12.3.179

Reyna, Valerie F., & Mills, B. A. (2014). Theoretically motivated interventions for reducing sexual risk taking in adolescence: A randomized controlled experiment applying fuzzy-trace theory. *Journal of Experimental Psychology: General*, *143*(4), 1627–1648. https://doi.org/10.1037/a0036717

Reyna, Valerie F, & Farley, F. (2006). Adolescent Implications for Theory , Practice , and Public Policy. *Psychological Science in the Public Interest*, *7*(1), 1–44.

Ruiz, J. G., Andrade, A. D., Garcia-Retamero, R., Anam, R., Rodriguez, R., & Sharit, J. (2013). Communicating global cardiovascular risk: Are icon arrays better than numerical estimates in improving understanding, recall and perception of risk? *Patient Education and Counseling*, *93*(3), 394–402. https://doi.org/10.1016/j.pec.2013.06.026

Schurr, P. H. (1987). Effects of Gain and Loss Decision Frames on Risky Purchase Negotiations. *Journal of Applied Psychology*, *72*(3), 351–358. https://doi.org/10.1037/0021-9010.72.3.351

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. https://doi.org/10.7326/0003-4819-127-11-199712010-00003

Sheridan, S. L., Halpern, D. J., Viera, A. J., Berkman, N. D., Donahue, K. E., & Crotty, K.

(2011). Interventions for individuals with low health literacy: A systematic review. *Journal of Health Communication*, *16*(SUPPL. 3), 30–54. https://doi.org/10.1080/10810730.2011.604391

Stone, E. R., Yates, J. F., & Parker, A. M. (1994). Risk communication: Absolute versus relative expressions of low-probability risks. *Organizational Behavior and Human Decision Processes*, Vol. 60, pp. 387–408. https://doi.org/10.1006/obhd.1994.1091

Sullivan, H. W., O'Donoghue, A. C., Lynch, M., Johnson, M., Davis, C., Amoozegar, J. B., & Rupert, D. J. (2021). Visual images of prescription drug benefits in direct-to-consumer television advertisements. *Patient Education and Counseling*, *104*(9), 2240–2249. https://doi.org/10.1016/j.pec.2021.02.024

Thomas, M., & Kyung, E. J. (2019). Slider Scale or Text Box: How Response Format Shapes Responses. *Journal of Consumer Research*, *45*(6), 1274–1293. https://doi.org/10.1093/jcr/ucy057

Thompson, V. A., Newman, I. R., Campbell, J. I. D., Kish-greer, C., Quartararo, G., & Spock, T. (2021). *Journal of Applied Research in Memory and Cognition Reasoning = representation + process : Common Ground for Fuzzy Trace and Dual Process Theories*. *10*, 532–536. https://doi.org/10.1016/j.jarmac.2021.10.008

Toçi, E. (2013). Health Literacy and Socioeconomic Characteristics among Older People in Transitional Kosovo. *British Journal of Medicine and Medical Research*, *3*(4), 1646–1658. https://doi.org/10.9734/bjmmr/2013/3972

Turpin, M. H., Meyers, E. A., Walker, A. C., Białek, M., Stolz, J. A., & Fugelsang, J. A. (2020). The environmental malleability of base-rate neglect. *Psychonomic Bulletin and Review*, *27*(2), 385–391. https://doi.org/10.3758/s13423-020-01710-1

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1007/978-1-4613-2391-4_2

van Weert, J. C. M., Alblas, M. C., van Dijk, L., & Jansen, J. (2021). Preference for and understanding of graphs presenting health risk information. The role of age, health literacy, numeracy and graph literacy. *Patient Education and Counseling*, *104*(1), 109–117. https://doi.org/10.1016/j.pec.2020.06.031

White, C. M., Gummerum, M., & Hanoch, Y. (2015). Adolescents' and Young Adults' Online Risk Taking: The Role of Gist and Verbatim Representations. *Risk Analysis*, *35*(8), 1407–1422. https://doi.org/10.1111/risa.12369

Wilhelms, E. A., Fraenkel, L., & Reyna, V. F. (2018). Effects of probabilities, adverse outcomes, and status quo on perceived riskiness of medications: Testing explanatory hypotheses concerning gist, worry, and numeracy. *Applied Cognitive Psychology*, *32*(6), 714–726. https://doi.org/10.1002/acp.3448

Wilhelms, E. A., & Reyna, V. F. (2013). *Fuzzy Trace Theory and Medical Decisions by Minors : Differences in Reasoning between Adolescents and Adults*. 268–282. https://doi.org/10.1093/jmp/jht018

Witteman, C. L. M., & Tollenaar, M. S. (2012). Remembering and diagnosing clients: Does experience matter? *Memory*, *20*(3), 266–276. https://doi.org/10.1080/09658211.2012.654799

Wolfe, C. R. (2021). Fuzzy-Trace Theory and the Battle for the Gist in the Public Mind. *Journal of Applied Research in Memory and Cognition*, *10*(4), 527–531. https://doi.org/10.1016/j.jarmac.2021.10.004

Wolfe, C. R., Dandignac, M., & Reyna, V. F. (2019). A theoretically motivated method for automatically evaluating texts for gist inferences. *Behavior Research Methods*, *51*(6), 2419–2437. https://doi.org/10.3758/s13428-019-01284-4

Wolfe, C. R., & Reyna, V. F. (2010). Semantic coherence and fallacies in estimating joint probabilities. *Journal of Behavioral Decision Making*, *23*(2), 203–223. https://doi.org/10.1002/bdm.650

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. https://doi.org/10.1177/0272989X13511706

APPENDIX A

A.1 Informed Consent Form

Bu araştırma, Kadir Has Üniversitesi Psikoloji Bölümü Öğretim Üyesi Dr. Öğr. Üy. Ahu Gökçe'nin danışmanlığında yüksek lisans öğrencisi Deniz Özdemir'in tez çalışması kapsamında yürütülmektedir.

Çalışmanın amacı bireylerin risk algısını incelenmektir. Sorulara verdiğiniz yanıtlar yalnızca bilimsel araştırma kapsamında kullanılacak ve araştırma ekibi dişinda kişilerle paylaşılmayacaktır. gizli tutulacaktır. Çalışmada kimliğinize yönelik herhangi bir veri toplanmayacaktır.
Araştırma yaklaşık 25 dakika sürmektedir. Herhangi bir sebepten ötürü çalışmadan istediğiniz zaman çekilme hakkına sahipsiniz.

Araştırmaya katılmak istiyorsanız lütfen aşağıdaki "kabul ediyorum" seçeneğini tıklayınız ve bir sonraki sayfaya geçiniz.
"Kabul ediyorum" seçeneğini tıklayarak bu onam formunu okuduğunuzu, anladığınızı ve araştırmaya katılmayı gönüllü olarak kabul ettiğinizi belirtmiş oluyorsunuz.

Onam formunu okudum, anladım. Araştırmaya katılmayı:

Kabul Ediyorum □
Kabul Etmiyorum □

A.2 Scenarios used in Experiment 1A

800-100 (Risk)

Bir otomotiv firması, yeni bir arabayı piyasaya çıkarmayı amaçlamaktadır ve bu modelde teknik özellikleri farklı 2 farklı seri araba üretilmiştir. Toplamda X serisinden 800 adet, Y serisinden 100 adet araba üretilmiştir. Satışa sunulan bu arabalarda 6 ay içinde X serisinde 16 arabada, Y serisinde 10 arabada arıza saptanıp, arızalanan arabalar üretici tarafından geri çağrılmıştır.

100-800 (Non-Risk)

Bir otomotiv firması, yeni bir arabayı piyasaya çıkarmayı amaçlamaktadır ve bu modelde teknik özellikleri farklı 2 farklı seri araba üretilmiştir. Toplamda T serisinden 100 adet, L serisinden 800 adet araba üretilmiştir. Satışa sunulan bu arabalarda 6 ay içinde T serisinde 2 arabada, L serisinde 80 arabada arıza saptanıp, arızalanan arabalar üretici tarafından geri çağrılmıştır.

900-150 (Risk)

Bir otomotiv firması, yeni bir arabayı piyasaya çıkarmayı amaçlamaktadır ve bu modelde teknik özellikleri farklı 2 farklı seri araba üretilmiştir. Toplamda M serisinden 900 adet, K serisinden 150 adet araba üretilmiştir. Satışa sunulan bu arabalarda 6 ay içinde M serisinde 18 arabada, K serisinde 15 arabada arıza saptanıp, arızalanan arabalar geri çağrılmıştır.

150-900 (Non-Risk)

Bir otomotiv firması, yeni bir arabayı piyasaya çıkarmayı amaçlamaktadır ve bu modelde teknik özellikleri farklı 2 farklı seri araba üretilmiştir. Toplamda G serisinden 150 adet, N serisinden 900 adet araba üretilmiştir. Satışa sunulan bu arabalarda 6 ay içinde G serisinde 3 arabada, N serisinde 90 arabada arıza saptanıp, arızalanan arabalar geri çağrılmıştır.

1200-200 (Risk)

Bir otomotiv firması, yeni bir arabayı piyasaya çıkarmayı amaçlamaktadır ve bu modelde teknik özellikleri farklı 2 farklı seri araba üretilmiştir. Toplamda B serisinden 1200 adet, D serisinden 200 adet araba üretilmiştir. Satışa sunulan bu arabalarda 6 ay içinde B serisinde 24 arabada, D serisinde 20 arabada arıza saptanıp, arızalanan arabalar geri çağrılmıştır.

200-1200 (Non-Risk)

Bir otomotiv firması, yeni bir arabayı piyasaya çıkarmayı amaçlamaktadır ve bu modelde teknik özellikleri farklı 2 farklı seri araba üretilmiştir. Toplamda Z serisinden 200 adet, P serisinden 1200 adet araba üretilmiştir. Satışa sunulan bu arabalarda 6 ay içinde Z serisinde 4 arabada, P serisinde 120 arabada arıza saptanıp, arızalanan arabalar üretici tarafundan geri çağrılmıştır.

A.3 Scenarios and Instructions used in Experiment 1B

800-100 (Risk)

Bir biyokimya firması tarafından geliştirilen ve kalp krizine karşı koruma sağlayan Clapizole adında bir ilacın ileri test aşamalarında 900 katılımcı yer almıştır. Deney grubunda olan 800 kişiye gerçek ilaç, kontrol grubunda olan 100 kişiye ise etkinliği olmayan boş (plasebo) ilaç verilmiştir. 6 aylık bir takip sonucunda katılımcıların kalp krizi geçirip geçirmedikleri kontrol edilmiştir. Sonuçlara göre gerçek ilacı alan 16 kişi, kontrol grubunda olan 10 kişi hastalığa yakalanmıştır.

100-800 (Non-Risk)

Bir biyokimya firması tarafından geliştirilen ve kalp krizine karşı koruma sağlayan Tacxolol adında bir ilacın ileri test aşamalarında 900 katılımcı yer almıştır. Deney grubunda olan 100 kişiye gerçek ilaç, kontrol grubunda olan 800 kişiye ise etkinliği olmayan boş (plasebo) ilaç verilmiştir. 6 aylık bir takip sonucunda katılımcıların kalp krizi geçirip geçirmedikleri kontrol edilmiştir. Sonuçlara göre gerçek ilacı alan 2 kişi, kontrol grubunda olan 80 kişi hastalığa yakalanmıştır.

900-150 (Risk)

Bir biyokimya firması tarafından geliştirilen ve kalp krizine karşı koruma sağlayan Olaxacol adında bir ilacın ileri test aşamalarında 1050 katılımcı yer almıştır. Deney grubunda olan 900 kişiye gerçek ilaç, kontrol grubunda olan 150 kişiye ise etkinliği olmayan boş (plasebo) ilaç verilmiştir. 6 aylık bir takip sonucunda katılımcıların kalp krizi geçirip geçirmedikleri kontrol edilmiştir. Sonuçlara göre gerçek ilacı alan 18 kişi, kontrol grubunda olan 15 kişi hastalığa yakalanmıştır.

150-900 (Non-Risk)

Bir biyokimya firması tarafından geliştirilen ve kalp krizine karşı koruma sağlayan Clatolol adında bir ilacın ileri test aşamalarında 1050 katılımcı yer almıştır. Deney grubunda olan 150 kişiye gerçek ilaç, kontrol grubunda olan 900 kişiye ise etkinliği olmayan boş (plasebo) ilaç verilmiştir. 6 aylık bir takip sonucunda katılımcıların kalp krizi geçirip geçirmedikleri kontrol edilmiştir. Sonuçlara göre gerçek ilacı alan 3 kişi, kontrol grubunda olan 90 kişi hastalığa yakalanmıştır.

1200-200 (Risk)

Bir biyokimya firması tarafından geliştirilen ve kalp krizine karşı koruma sağlayan Omdinol adında bir ilacın ileri test aşamalarında 1400 katılımcı yer almıştır. Deney grubunda olan 1200 kişiye gerçek ilaç, kontrol grubunda olan 200 kişiye ise etkinliği olmayan boş (plasebo) ilaç verilmiştir. 6 aylık bir takip sonucunda katılımcıların kalp krizi geçirip geçirmedikleri kontrol edilmiştir. Sonuçlara göre gerçek ilacı alan 24 kişi, kontrol grubunda olan 20 kişi hastalığa yakalanmıştır.

200-1200 (Non-Risk)

Bir biyokimya firması tarafından geliştirilen ve kalp krizine karşı koruma sağlayan Topiritol adında bir ilacın ileri test aşamalarında 1400 katılımcı yer almıştır. Deney grubunda olan 1200 kişiye gerçek ilaç, kontrol grubunda olan 200 kişiye ise etkinliği olmayan boş (plasebo) ilaç verilmiştir. 6 aylık bir takip sonucunda katılımcıların kalp krizi geçirip geçirmedikleri kontrol edilmiştir. Sonuçlara göre gerçek ilacı alan 4 kişi, kontrol grubunda olan 120 kişi hastalığa yakalanmıştır.

A.4 Numeracy Scale

Aşağıdaki sorularda size doğru gelen şıkkı işaretleyiniz. Bazı sorularda seçenek yok, bu sorularda cevabınızı boş bırakılan kutucuğa **rakam** olarak giriniz.

Questions		Correct Answers
1	Aşağıdaki sayı aralıklarından hangisi, bir hastalığa yakalanma riskini en fazla temsil eder? A)100'de 1 B) 1000'de 1 C) 10'da 1	C)10'da 1
2	Aşağıdaki yüzdelerden hangisi, bir hastalığa yakalanma riskini en fazla temsil eder? A)%1 B) % 10 C) % 5	B) %10
3	3. Bir hastalığa yakalanma olasılığı % 10 ise, 100 kişiden kaç kişinin bu hastalığa yakalanması beklenir?	10
4	Bir hastalığa yakalanma olasılığı % 10 ise, 1000 kişiden kaç kişinin bu hastalığa yakalanması beklenir?	100
5	Bir hastalığa yakalanma olasılığı 100 üzerinden 20 ise, bu hastalığa yakalanma olasılığı % ____ 'dir.	20
6	A kişisinin hastalığa yakalanma riski on yıl içinde % 1'dir. B kişisinin hastalığa yakalanma riski A kişisinin hastalığa yakalanma riskinin iki katı ise, B kişisinin hastalığa yakalanma riski yüzde kaçtır?	%2
7	A kişisinin hastalığa yakalanma riski on yıl içinde 100'de 1'dir. B kişisinin hastalığa yakalanma riski A kişisinin hastalığa yakalanma riskinin iki katı ise, B kişisinin hastalığa yakalanma riski kaçta kaçtır?	%2
8	Piyangoda 10 dolar ödül kazanma şansı % 1'dir. 1000 kişinin her biri tek bir piyango bileti aldığında, kaç kişinin 10 dolar ödülü kazanacağına dair tahmininiz nedir?	10
9	Hilesi olmayan, 1'den 6'ya sayıların yer aldığı bir zarı 1000 defa attığınızı düşünün. Sizce 1000 atışın kaç tanesinde zar çift sayı (2/4/6) getirir?	500
10	Viral enfeksiyon kapma olasılığı 0.0005'tir. 10.000 kişiden kaç tanesine virüs bulaşması beklenir?	5
11	Çekilişte bir araba kazanma şansı 1000'de 1'dir. Çekilişte biletlerin yüzde kaçı bir araba kazandırır?	%0.1

A.5 Turkish Version of Health Literacy Scale

Aşağıdaki sorular sağlıkla ilgili bilgileri nasıl değerlendirdiğinizle alakalı kişisel düşüncelerinizi ölçmeye yöneliktir. Belirtilen cümleyle ilgili beş adet seçenekten kendinize en yakın seçeneğini işaretleyiniz.

1. Hastalıklar hakkında bilgileri bulabiliyor musunuz?

2. Tedaviler hakkında bilgileniyor musunuz?

3. Sigara içme, şişmanlık gibi sağlık riskleri hakkında bilgileri bulabiliyor musunuz?

4. Nasıl sağlıklı kalınacağı hakkında bilgileri bulabiliyor musunuz?

5. Sağlıklı yiyecekler ve nasıl formda kalınacağı hakkında bilgileri elde edebiliyor musunuz?

6. İlaç kutularında bulunan açıklayıcı bilgileri anlayabiliyor musunuz?

7. Tıbbi reçeteleri anlayabiliyor musunuz?

8. Eczanelerde, hastanelerde ya da doktor bulunan sağlığa zararlı davranışlar hakkında bilgileri sağlayan broşürleri okuyabiliyor musunuz?

9. Sigara İçmek, uyuşturucu kullanmak, içkili araba kullanmak vb. gibi tehlikeli davranışlar hakkındaki bilgileri anlayabiliyor musunuz?

10. Besin etiketlerinin içeriğini anlayabiliyor musunuz?

11. Sağlıklı yaşam biçiminin önemini anlayabiliyor musunuz?

12. Ev, okul, işyeri ya da mahallede sağlıklı çevrenin önemini anlayabiliyor musunuz?

13. Doktorunuzla ya da eczacınızla tıbbi bilgileri tartışabiliyor musunuz?

14. Tedavi seçeneklerinin yan etkilerini ya da yararlarını düşünebiliyor musunuz?

15. Tıbbi önerilerden hangisinin sizin için en iyisi olduğuna karar verebiliyor musunuz?

16. Sağlığınıza zararlı davranışlarınızı belirleyebiliyor musunuz?

17. Diğer İnsanların yaptığı sağlığa zararlı davranışlardan ders alabiliyor musunuz?

18. Sağlık personeli, arkadaşlarınız, aileniz ya da radyo, gazete, televizyon gibi kaynaklardan edindiğiniz sağlığa zararlı davranışlarla ilgili bilgileri dikkatli biçimde değerlendirebiliyor musunuz?

19. Sağlıkla ilgili alışkanlıklarınızı değerlendirebiliyor musunuz?

20. Sağlıklı beslenme ya da spor gibi sağlıklı seçimlerin etkilerini ve yararlarını düşünebiliyor musunuz?

21. Doktor, hemşire ya da eczacının size verdiği önerilere uyabiliyor musunuz?

22. Aşı yaptırma, bir tarama programında yer alma, güvenli araba kullanma gibi sağlık personellerinin size verdiği önerilere uyabiliyor musunuz?

23. Eğer isterseniz sağlığa zararlı alışkanlıklarınızı değiştirebiliyor musunuz?

24. Sağlıklı ürünlere (doğal besinler, zararsız kimyasallar gibi) ulaşabiliyor musunuz?

25. Sağlıkla ilgili bilgileri sizin yararınıza olacak şekilde kullanabiliyor musunuz?

(Yapamayacak durumdayım=1; Çok zorluk çekiyorum=2; Biraz zorluk çekiyorum=3; Az zorluk çekiyorum=4; Hiç zorluk çekmiyorum=5)

A.6 CVD Risk and Subjective Risk of CVD

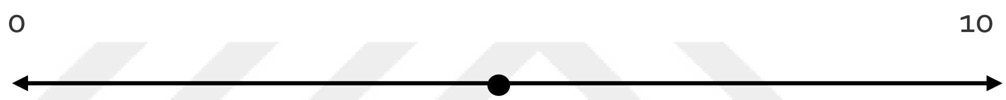
Bu soru kalp-damar hastalıkları hakkında aile öykünüzü öğrenmek amacıyla sorulmuştur.

1. Geçmişte size veya ailenizden birine kalp damar hastalığı tanısı konuldu mu?

   Evet ☐

   Hayır ☐

2. Aşağıdaki ölçekte şimdi ve gelecekte kalp-damar hastalıklarına karşı ne kadar risk altında olduğunuzu hissettiğinizi belirtiniz. (0=risk altında değilim, 5=orta derecede risk altındayım, 10=yüksek derecede risk altındayım)



A.7 Dependent Variables

Risk Reduction Estimates

In experiment 1A, two questions were asked and accuracy was calculated accordingly. Correct answer was always 20 and 100 (Q1 and Q2, respectively) for experiment 1A.

Q1: X serisinden 1000 araç üretilirse, kaç araç arızalanır?

Q2: Y serisinden 1000 araç üretilirse, kaç araç arızalanır?

In experiment 1B, two questions were asked and accuracy was calculated accordingly. Three types of scales were used.

Correct answer was 10 and 50 for group sizes of 800-100 and 100-800;

Q1: Bu ilacı alan 500 kişilik bir grupta yaklaşık kaç kişi kalp krizi geçirebilir?

Q2: Bu ilacı almayan 500 kişilik bir grupta yaklaşık kaç kişi kalp krizi geçirebilir?

Correct answers were 20 and 100 for group sizes of 900-150 and 150-900;

Q1: Bu ilacı alan 1000 kişilik bir grupta yaklaşık kaç kişi kalp krizi geçirebilir?

Q2: Bu ilacı almayan 1000 kişilik bir grupta yaklaşık kaç kişi kalp krizi geçirebilir?

Correct answers were 40 and 200 for group sizes of 1200-200 and 200-1200;

Q1: Bu ilacı alan 2000 kişilik bir grupta yaklaşık kaç kişi kalp krizi geçirebilir?

Q2: Bu ilacı almayan 2000 kişilik bir grupta yaklaşık kaç kişi kalp krizi geçirebilir?

Confidence

Q: Verdiğiniz cevapların doğruluğundan ne kadar eminsiniz?

Willingness

Experiment 1A: (willingness; Q1 = better, Q2=worse)

Q1: X serisindeki aracı kullanmak için ne kadar istekli olurdunuz?

Q2: Y serisindeki aracı kullanmak için ne kadar istekli olurdunuz?

Experiment 1B:

Q: Kalp krizi riskini azaltmak için bu ilacı almaya ne kadar istekli olurdunuz?

CURRICULUM VITAE

Personal Information

Name Surname : Deniz Özdemir

Education

Undergraduate Education : Kadir Has University, Psychology (2013-2018)

Graduate Education : Kadir Has University, Psychological Sciences (2019-2022)

Foreign Languages : English