IN PRESS

JOURNAL OF BEHAVIORAL DECISION MAKING

How much risk can you stomach?

Individual differences in the tolerance of perceived risk across gender and risk domain

Jonathan J. Rolison

Department of Psychology, University of Essex, UK

Jeordie Shenton

Department of Psychology, University of Essex, UK

Correspondence to: Jonathan J Rolison, Department of Psychology, University of Essex, Essex, CO4 3SG, UK; Tel: +44 (0)28 9097 5653, Email: jrolison@essex.ac.uk

Key words

Risk-taking, Risk attitude, Risk perception, DOSPERT, risk domain, gender differences

Word count (main text): 10,687

Abstract

Research within the psychological risk-return framework, namely using the Domain-Specific Risk-Taking (DOSPERT) scale, has led to a conclusion that risk attitude—measured as an individual's sensitivity to the risk they perceive—is stable across people (e.g., gender) and domains (e.g., recreational, social, financial, health). Risk-taking differences across gender and domain have been interpreted in terms of differences in the magnitude of risk perceived (and expected benefit). Yet, the DOSPERT scale items, contrived by researchers, rather than decision-makers themselves, may have failed to detect differences in perceived risk attitude by failing to adequately represent all combinations of risks and benefits across gender and domains. In Study 1, participants generated their own examples of activities, which we selected among in Studies 2 and 3 to construct a new scale representing various levels of perceived risk and expected benefit. Our findings reveal that women are more sensitive than men to risk they perceive (i.e., are less tolerant of risk) in the recreational, social, and financial domains, but not the health domain. Risk attitude also differed across domains, with participants tolerating more risk in some domains than in others. We conclude that gender and domain differences in risk-taking stem partly from gender and domain differences in people's sensitivity to perceived risks. Our findings have theoretical implications for the psychological risk-return framework and bridge with other theoretical approaches, such as the expected utility framework. Our studies also provide a new scale for assessing differences in attitudes toward risk that overcomes shortcomings of existing scales.

How much risk can you stomach?

Individual differences in the tolerance of perceived risk across gender and risk domain

Why do some people wilfully engage in risky activities—going white-water rafting, taking illicit drugs, or gambling at a casino—whereas others shy away from risk? One possible explanation for individual differences in risk-taking is differences in the magnitude of risk that people perceive: A person may be more (or less) willing than another to engage in an activity because they perceive the activity to be less (or more) risky. A second possible explanation for individual differences in risk-taking is differences in the attitudes that people possess toward risk: A person may be more (or less) willing than another to engage in an activity because they are more (or less) tolerant of risk, despite perceiving the same magnitude of risk. Do apparent differences in risk-taking across people and contexts stem from both or either differences in the perceived magnitude of risk (i.e., risk perception) or differences in attitudes toward risk (i.e., risk attitude)?

Psychological risk-return models conceptualise risk-taking as a trade-off between the perceived riskiness and expected benefit of engaging in an activity (Weber, 1997; 1998; Weber, Blais, & Betz, 2002). An attractive feature of this approach is that differences in risk-taking behavior can be decomposed into differences in the magnitude of (a) perceived risk, (b) expected benefit, and (c) differences in risk attitude, which is characterized by an individual's degree of tolerance of risk to attain an expected benefit. In a linear regression model, in which risk-taking behavior is regressed onto risk perception and expected benefit ratings, risk attitude is represented by the coefficient for the effect of risk perception (Weber et al., 2002; Weber & Milliman, 1997). Accordingly, a large negative coefficient for risk perception would indicate a strong negative attitude toward risk, as for any given level of expected benefit a unit increase in perceived risk would be associated with a large decrease in willingness to engage in an activity. In comparison, a small negative coefficient for risk perception would indicate a high tolerance of risk, as a unit increase in perceived risk would be associated with a small decrease in willingness to engage in the activity.

A wealth of previous research within the psychological risk-return framework has led to the conclusion that differences in self-reported risk-taking behavior across people and contexts (e.g., risk domains) is attributable to differences in the perceived magnitude or risk (i.e., perceived riskiness of an activity), rather than to differences in risk attitude. For example, Weber et al. (2002) reported that women were more risk averse than men and that such gender differences in risk-taking were due to differences in risk perception and expected benefits, and not to differences in their attitude toward perceived risk (see also Figner & Weber, 2011; Hanoch, Johnson, & Wilke, 2006; Harris, Jenkins, & Glaser, 2006; Rolison, Hanoch, & Freund, in press; Rolison, Hanoch, Wood, & Pi-Ju, 2014; but see Zhang, Foster, & McKenna, in press). Cultural differences in risk-taking have also been attributed to differences in the perceived riskiness of activities, rather than to differences in attitudes toward risk (Weber & Hsee, 1998; 1999). Perceived risk attitude is also purported to be stable across risk domains, such that risk-taking behavior across domains differs as a function of domain differences in the perceived riskiness of activities (Hanoch, Johnson, & Wilke, 2006; Weber et al., 2002; Weber & Miliman, 1997). Thus, perceived risk attitude seemingly is stable across people (e.g., gender and culture) and contexts (e.g., risk domains).

Evidence for the stability of attitudes toward risk across people and domains has been delivered primarily by research on the Domain-Specific Risk-Taking (DOSPERT) scale, which was constructed within the psychological risk-return framework to measure individual differences in the trade-off between perceived risk and expected benefit across five risk domains, including the recreational, financial, health, social, and ethical domains (Weber et al., 2002). The DOSPERT scale contains three components: risk-taking behavior—a self-report of the likelihood of engaging in an activity; risk perception—a rating of the magnitude of risk perceived for an activity; and expected benefit—a rating of the magnitude of benefit expected for an activity. For each of a number of activities (e.g., 'trying out bungee jumping at least once'; the recreational domain), participants rate their likelihood of engagement, risk perception, and expected benefit on a Likert scale. The DOSPERT has been shown to have reasonable levels of internal consistency (Rolison et al., 2014; Weber et al., 2002), test-retest reliability (Weber et al., 2002), and to capture behavioral tendencies that are at least partly domain-specific in nature (Highhouse, Nye, Zhang, & Rada, 2017).

However, in the DOSPERT, the assessment of risk attitude may be thwarted by skewed behavior on the scale. To demonstrate, Figure 1 provides the ratings of 528 participants on the likelihood of engagement (i.e., risk-taking behavior) subscale of the DOSPERT in the recreational, social, financial, and health domains, reproduced from Rolison et al. (2014). As shown in Figure 1, participants' ratings in the recreational, financial, and health domains were positively skewed toward the low end of the scale (i.e., unlikely to engage). In fact, the lowest point on the scale (value = 1, "extremely unlikely") accounted for around half of participants' ratings in these three domains. Conversely, participants' ratings in the social domain instead showed a negative skew toward the high end of the scale (i.e., likely to engage). Rolison et al. (2014) did not measure perceived risk attitude in their study. However, Figure 1 indicates that when likelihood of engagement is regressed onto risk perception to measure perceived risk attitude, the range of likelihood of engagement that informs perceived risk attitude differs across domains. This is because some domains of the DOSPERT contain items that people are unlikely to engage in (e.g., 'going down a ski run that is beyond your ability'; recreational domain) whereas other domains contain activities that people are likely to engage in (e.g., 'admitting that your tastes are different from those of a friend'; social domain). Therefore, the effect of a unit increase in risk perception on likelihood of engagement in one domain may not be comparable with the effect of a unit increase in risk perception in another domain when domains differ greatly in the riskiness of their activities.

A second related issue is that higher risk perceptions are typically associated with lower expectations of benefit (Alhakami & Slovic, 1994; Slovic, 1997; Weber et al., 2002). Consequently, some combinations of risk perception and expected benefit are more frequent than others. For instance, activities are more likely to receive a high risk perception rating and a low rating of expected benefit than they are to receive a low risk perception rating and a low rating of expected benefit. This behavioral tendency impedes a reliable assessment of risk attitude within the psychological risk-return framework when not all combinations of ranges of risk perception and expected benefit are adequately represented in a set of activities. In the DOSPERT, women report that they expect fewer benefits from engaging in activities than do men (Hanoch et al., 2006; Weber et al., 2002) and expected benefit ratings differ across risk domains (Weber et al., 2002). When likelihood of engagement is regressed onto risk perception and expected benefit ratings in a linear regression analysis, it is tacitly assumed that expected benefit ratings are constant across individuals and domains. Yet, combinations of levels of risk perception and expected benefit are likely to differ across individuals and domains. Therefore, in the assessment of risk attitude in the DOSPERT, the effect of a unit increase in risk perception on likelihood of engagement is likely to be assessed at differing levels of expected benefit across domains and individuals (e.g., gender). In the current studies, we develop a new domain-specific risk-taking scale that samples items at low, medium, and high levels of risk perception and expected benefits to enable the assessment of individual differences and domain differences in perceived risk attitude.

While research using the DOSPERT has indicated that attitudes toward risk are stable across people and domains, studies of risky choice behavior have reported individual differences in risk attitude. Within the expected utility framework, a common method for assessing risk attitude has been to compare choice behavior for a lottery over varying amounts of a sure outcome. Based on an individual's choice behavior, their risk attitude is determined by the shape of their utility function. Accordingly, choice of a lottery over a sure amount of equal expected value is indicative of a risk seeking policy, whereas choice of the equal sure amount is diagnostic of a risk averse policy. In contrast with the psychological risk-return framework, risk perception and attitude toward perceived risk are not distinguished within the expected utility framework. Regarding individual differences, Sapienza, Zingales, and Maestripieri (2009) found that compared with women, men more frequently accepted a monetary gamble over equivalent guaranteed amounts, suggesting greater risk-tolerance among men than women (see also Charness & Gneezy, 2012; Levin, Snyder, & Chapman, 1988; Powell & Ansic, 1997; Rosen, Tsai, & Downs, 2003). In Charness and Gneezy (2012), the authors concluded based on a review of 15 studies that men are more risk seeking than women for risky financial investment tasks. However, with few exceptions (e.g., Rosen et al., 2003), studies within the risky choice framework have focussed on decisions about monetary outcomes. Thus, it remains unclear whether risk attitude differs across risk domains, such as those captured by the DOSPERT, and whether possible gender differences in risk attitude are stable (or differ) across risk domains.

In sum, research within the psychological risk-return framework has led to a conclusion that perceived risk attitude—i.e., the tolerance of perceived risk to attain a benefit—is stable across people and domains. Yet, research conducted within the expected utility framework, with a focus on choice behavior in the financial domain, suggests that risk attitude differs across people (e.g., gender). Our aim was to test the stability of risk attitude across gender and risk domain, reconciling the issues discussed earlier regarding the assessment of risk attitude in the DOSPERT.

To do so, in Study 1, we asked participants to generate their own examples of real-life activities in each of four domains (recreational, financial, social, health), producing a large sample of real-life activities and circumventing the need to contrive a set of items for each risk domain. To ensure a broad range of activities according to perceived risk and expected benefit, some participants were asked to freely generate activities (control condition), others were asked to generate examples of very risky activities (high-risk condition), and others were asked to generate examples of highly beneficial activities (high-benefit condition). Participants then rated their risk perceptions and expected benefits for the activities they generated. In Study 2, we selected among the activities generated by participants in Study 1, selecting items for each domain that represented various combinations of levels of risk perception and expected benefit (e.g., low risk perception and low expected benefit, low risk perception and high expected benefit etc.). Using our newly developed scale, we conducted a multilevel regression analysis to test for differences in risk attitude across gender and risk domain. In Study 3, we revised the scale items to improve the scale reliability and structure.

As previous research has shown risk perception and expected benefit to be negatively correlated (e.g., Alkakami & Slovic, 1994), we hypothesized that in Study 1 higher perceived risk would be associated with lower expected benefit. Studies of risky choice behavior within the expected utility framework have indicated that men are more tolerant than women of financial risks (e.g., Charness & Gneezy, 2012). Thus, we hypothesized that in Studies 2 and 3 men would be more tolerant of financial risk than women, indicated by lower sensitivity to perceived risk among male participants. However, as studies within the expected utility framework have focussed on decisions about monetary outcomes, it is unclear whether gender differences in risk attitude will be limited to the financial domain or will be stable across risk domains. In Studies 2 and 3, we further explored whether risk attitude differs across risk domains as previous studies using the DOSPERT may have failed to capture possible domain differences in risk attitude due to the problems we have identified with the scale. While we had no specific predictions about whether gender differences in risk attitude would differ across risk domains, Studies 2 and 3 further permitted us to explore this possibility.

Study 1

In Study 1, participants generated their own examples of real-life activities in each of four domains (recreational, financial, social, health) and rated each of the items they generated according to their risk perception and expected benefits.

Method

Participants

Two hundred seventy participants residing in the US were recruited online using Amazon's Mechanical Turk. Studies have shown that the Mechanical Turk platform is comparable with other methods of testing (e.g., face-to-face behavioral testing) and provides more socio-economically diverse sample characteristics than other recruitment methods (Casler, Bickel, & Hackett, 2013; Gibson, Piantadosi, & Fedorenko, 2011). Nineteen participants were excluded either because they provided nonsense written responses to the risky scenario items ($n = 9$; see Materials and Procedure) or because they were located outside of the US as determined by inspection of their computer IP address ($n = 10$). Our final sample included 251 participants (55% female, mean age = 36.50, $SD = 11.96$). Regarding education, 20 (8%) indicated high school as their highest educational attainment, 78 (31%) indicated some college, 120 (48%) indicated university, and 33 (13%) indicated postgraduate education. Regarding employment, the majority were in full- or part-time employment ($n = 194$; 77%).

Materials and procedure

Generating activities: For each of four life domains—recreational, social, health, and financial[1]—participants were asked to generate four examples of activities. In the control condition (*n* = 88), the instructions read:

"*Please describe in one sentence 4 [**recreational, financial, health, social**] activities or behaviors that you could engage in. The activities or behaviors you generate could be ones that you already engage in or ones that you do not currently engage in.*"

In the high-benefit condition (*n* = 84), participants were explicitly instructed to generate examples of activities that they perceived to be highly beneficial for them. The instructions read:

"*Please describe in one sentence 4 [**recreational, financial, health, social**] activities or behaviors that would be **highly beneficial** for you to engage in. The activities or behaviors you generate could be ones that you already engage in or ones that you do not currently engage in.*"

In the high-risk condition (*n* = 84), participants were explicitly instructed to generate examples of activities that they perceived to be very risky for them. Participants were asked to generate activities that were risky for them rather than for people in general as we expected that they would be better able to assess risks that are personal to them. The instructions read:

"*Please describe in one sentence 4 [**recreational, financial, health, social**] activities or behaviors that would be **very risky** for you to engage in. The activities or behaviors you generate could be ones that you already engage in or ones that you do not currently engage in.*"

We did not ask participants to generate examples of low-risk activities because we expected that participants in the control condition would generate many low-risk examples as

---

[1] We did not include an ethical domain, which is one of the five domains of the DOSPERT (Weber et al., 2002). First, we reasoned that it was inappropriate to ask participants to generate multiple examples of unethical behavior. Second, unlike behavior in other domains, unethical behavior is likely to be governed by moral and ethical values rather than a trade-off of perceived risks and expected benefits (Weber et al., 2002).

they were not told to consider risky activities. We also did not ask participants to generate examples of low-benefit activities as participants in the control condition were asked to generate examples of activities that they do not currently engage in as well as ones that they already engage in, encouraging a broad range of activities according to their expected benefits. Moreover, as risky activities are often perceived to yield fewer benefits (e.g., Alhakami & Slovic, 1994), we expected that the high-risk condition would generate many examples of low-benefit activities.

Once participants had described four activities for a domain, they moved onto the next domain, presented on a separate page. The four domains were presented in a randomly generated order for each participant.

Evaluating activities: Participants were then presented the 16 activities they previously generated and were asked to evaluate their risk perceptions and expected benefits. For risk perception, participants were told:

"People often see some risk in situations that contain uncertainty about what the outcome or consequence will be and for which there is a possibility of negative consequences. However, riskiness is a very personal and intuitive notion, and we are interested in ***your gut level assessment of risk****.*

Below, are the activities or behaviors that you previously generated. Please indicate how ***risky you perceive*** *each* ***for you personally*** *if you were to engage in the activity or behavior."*

Participants provided their ratings on a 10-point scale, ranging 1 ("Not at all risky") to 10 ("Extremely risky").

For expected benefits, participants were told:

Below, are the activities or behaviors that you previously generated. Please indicate the ***benefits*** *you would obtain from engaging in each one."*

Participants provided their ratings on a 10-point scale, ranging 1 (“No benefits at all”) to 10 (“Great benefits”). The instructions and rating scales for evaluating risk perception and expected benefits were based on those used in the DOSPERT scale (Weber et al., 2002).

The 16 activities were presented in a randomly generated order within each section for each participant. Participants were randomly assigned to first complete either the risk perception or expected benefit section.

Results

Figure 2 provides participants’ mean expected benefits and risk perceptions for the activities they generated under control, high-benefit, and high-risk conditions. In the control condition, participants were unguided in their generation of activities—i.e., they were not asked to generate highly beneficial or highly risky activities—enabling us to assess the naturally occurring relationships among expected benefits and risk perceptions for activities. In the control condition, the inter-correlations were high among the four risk domains, except for the financial domain (Table 1). Specifically, participants who expected greater benefit or perceived higher risk in one domain, did so also in other domains (Table 1). The financial domain is an exception as participants who perceived higher risk in recreational and social domains did not perceive greater risk also in the financial domain. Within the risk domains, expecting greater benefit was associated with perceiving lower risk in the recreational ($r(88) = -.32$, $p = .002$), social ($r(88) = -.27$, $p = .011$), financial ($r(88) = -.33$, $p = .002$), and health ($r(88) = -.33$, $p = .002$) domains.

When asked to generate examples of very risky activities, do participants perceive them to be more (or less) beneficial? As shown in Figure 2, participants expected greatest benefits in the high-benefit condition ($M = 7.97$), followed by the control ($M = 6.72$), and high-risk ($M = 3.13$) conditions. A two-way mixed analysis of variance (ANOVA) on expected benefits, including condition (control, high-benefit, high-risk) and risk domain (recreational, social, financial, health) as factors, confirmed a significant effect of condition ($F(2,248) = 266.00$, $p < .001$, $\eta^2 = .68$). Planned pairwise comparisons confirmed that expected benefits were significantly greater in the high-benefit condition ($p < .001$) and lower in the high-risk condition ($p < .001$) compared to the control condition. There was also an effect of risk domain ($F(3,744) = 35.33$, $p < .001$, $\eta^2 = .13$), whereby benefits were perceived to be greatest in the health domain ($M = 6.54$), followed by the recreational ($M = 6.20$), social ($M = 5.57$), and financial ($M = 5.45$) domains. However, condition interacted with domain ($F(6,744) = 26.75$, $p < .001$, $\eta^2 = .18$). As shown in Figure 2, the high-benefit and high-risk conditions appeared to dampen the effect of domain on expected benefits. Follow-up ANOVAs confirmed that the effect of domain was stronger in the control condition ($F(3,261) = 60.56$, $p < .001$, $\eta^2 = .41$) than in the high-benefit ($F(3,249) = 19.86$, $p < .001$, $\eta^2 = .19$) and high-risk ($F(3,234) = 9.24$, $p < .001$, $\eta^2 = .11$) conditions. In Figure 2, the mean expected benefit ratings are close to the high end of the rating scale in the high-benefit condition and are close to the low end of the scale in the high-risk condition. Hence, the limited range of the scale may have caused the dampened effects of domain in these conditions relative to the control condition.

When asked to generate examples of highly beneficial activities, do participants perceive them to be more (or less) risky? As shown in Figure 2, participants perceived slightly lower risk overall for activities in the high-benefit ($M = 3.02$) condition compared to the control ($M = 3.59$) condition and perceived far higher risk in the high-risk condition ($M = 7.88$). A two-way mixed ANOVA on risk perceptions, including condition and domain as factors, confirmed a significant effect of condition ($F(2,248) = 320.51$, $p < .001$, $\eta^2 = .72$). Planned comparisons confirmed that perceived risk was significantly higher in the high-risk condition

($p < .001$) and lower in the high-benefit condition ($p < .005$) compared to the control condition. There was also an effect of domain ($F(3,744) = 53.47$ $p < .001$, $\eta^2 = .18$), in which risk was perceived to be greatest in the financial domain ($M = 5.79$), followed by the recreational ($M = 4.65$), social ($M = 4.65$), and health ($M = 4.23$) domains. However, condition interacted with domain ($F(6,744) = 34.48$, $p < .001$, $\eta^2 = .22$). As we observed for expected benefits, the high-benefit and high-risk conditions appeared to dampen the effect of domain on risk perceptions (Figure 2). Follow-up ANOVAs confirmed that the effect of domain was stronger in the control condition ($F(3,261) = 82.31$, $p < .001$, $\eta^2 = .49$) than in the high-benefit ($F(3,249) = 24.46$, $p < .001$, $\eta^2 = .23$) and high-risk ($F(3,234) = 4.76$, $p = .003$, $\eta^2 = .06$) conditions. Independent samples *t*-tests further showed that participants indicated a significantly lower risk perception in the high-benefit condition than in the control condition only in the financial domain ($t(170) = 4.91$, $p < .001$), and not in the recreational ($t(170) = 0.33$, $p = .741$), social ($t(170) = 0.62$, $p = .537$), or health ($t(170) = 1.53$, $p = .127$) domains. The mean risk perception ratings were close to the high end of the rating scale in the high-risk condition, which may have caused the dampened effects of domain relative to the control condition.

Summary

In sum, as predicted, when participants were asked to generate activities and then to rate them according to their expected benefits and risk perceptions, expecting more of the former was associated with perceiving less of the latter. Moreover, participants who were asked to generate examples of very risky activities rated the activities as less beneficial, and in the financial domain, participants who were asked to generate highly beneficial activities rated them as less risky.

Study 2

In Study 2, we selected among the activities generated by participants in Study 1 to create a new risk-taking scale that represents all combinations of low, medium, and high levels of risk perception and expected benefits. Using our newly developed scale, we tested for differences in perceived risk attitude across gender and risk domain.

Method

Participants

Two hundred seventy-one participants residing in the US were recruited online using Amazon's Mechanical Turk. Fifteen participants were excluded because they were located outside of the US as determined by their computer IP address. Our final sample included 256 participants (39% female, mean age = 35.61, *SD* = 11.90). Regarding education, 26 (10%) indicated high school as their highest educational attainment, 81 (32%) indicated some college, 119 (46%) indicated university, and 30 (12%) indicated postgraduate education. Regarding employment, the majority were in full- or part-time employment (*n* = 191; 75%).

Materials and procedure

We generated a 36-item scale using a sample of the activities generated by participants in Study 1.[2] To include activities that capture all combinations of expected benefit and risk perception (e.g., high-risk and low-benefit, low-risk and high-benefit etc.), we categorized the activities according to their expected benefit and risk perception ratings. Activities were categorized as low expected benefit (perceived risk) if they received a rating of 1 to 3, as medium expected benefit (perceived risk) if they received a rating of 4 to 7, and as high expected benefit (perceived risk) if they received a rating of 8 to 10. In other words, activities were categorized as either low, medium, or high expected benefit and perceived risk based on the ratings they received by participants in Study 1. Finally, for each risk domain, the most

[2] Items that were similar (e.g., "smoking", "smoking cigarettes", "smoking cigars") were combined under generic phrasing (i.e., "smoking"). To improve clarity, examples were included for some items (e.g., 'Unhealthy eating [e.g., eating a hamburger]') that had been generated by participants.

frequently generated activities were chosen to represent each of the nine possible combinations of the expected benefit and risk perception categories, generating a total of 36 items. Appendix A provides the 36 scale items, their category (i.e., low, medium, high) of expected benefit and risk perception, and the number of male and female participants who generated each item in Study 1.

For the 36-item scale, participants were asked to evaluate their expected benefits, risk perceptions, and likelihood of engagement for the items in three separate blocks. They provided their expected benefits and risk perceptions for the items in the same way as in Study 1. For likelihood of engagement, participants were asked:

"*For each of the following statements, please indicate the* ***likelihood*** *that you would engage in each activity or behavior if you were to find yourself in that situation."*

Participants provided their ratings on a 10-point scale, ranging 1 ("Extremely unlikely") to 10 ("Extremely likely").

The 36 activities were presented in a randomly generated order within each block for each participant. Participants were randomly assigned to first complete either the risk perceptions or expected benefits section, completing the likelihood rating section last.

Results

We first assessed whether the scale items that were selected from Study 1 for the low, medium, and high expected benefit categories received corresponding low, medium, and high expected benefit ratings by participants in Study 2, and whether the items selected for the low, medium, and high risk perception categories received corresponding low, medium, and high risk perception ratings. Briefly, this analysis confirmed that expected benefit ratings increased with the expected benefit category (i.e., low, medium, to high) and that risk perception ratings increased with the risk perception category (see Appendix B for full description).

Thus, as intended, our scale captured a range of levels of expected benefit and risk perception.

Scale reliability

Provided in Table 2 are the mean item-total correlations, which indicate the degree to which the individual scale items correlated with the total scores of each subscale in each domain. The mean item-total correlations were all above .30, which has been recommended as a minimum criterion value (Nunnally & Bernstein, 1994). However, within the subscales, some item-total correlations fell below .30. Four items in the recreational domain, one item in the financial domain, and one item in the health domain exhibited item-total correlations that were considerably below .3 for multiple subscales. Hence, these items may require replacement during scale refinement in Study 3. The Cronbach alpha scores, measuring scale internal consistency reliability, were above .70 for each of the subscales in the recreational, social, and financial domains, indicating acceptable levels of internal consistency (see Kline, 1999). The health domain, however, exhibited the poorest Cronbach alpha scores, which were low for all three subscales.

Inter-correlations of expected benefits and risk perceptions across domains

We assessed the inter-correlations for expected benefits and risk perceptions across the four risk domains (Table 3). The scale items in Study 2 were selected to capture all combinations of expected benefits and risk perceptions (e.g., high-risk and low-benefit, low-risk and high-benefit etc.) in order to match the domains according to their range of values. We examined whether domain-specificity, as indicated by low inter-correlations across domains, is reduced when the range of expected benefits and risk perceptions is matched across domains. The inter-correlations were higher across domains in comparison to Study 1 (Table 1). To compare the correlations in Study 1 and Study 2 we employed the method proposed by Cohen and Cohen (1983) for comparing correlations based on independent samples, which involves comparing z-scores following Fisher's r-to-z transformation (see Preacher, 2002). Sixteen of the 18 inter-correlations were significantly higher in Study 2 compared to Study 1 (Table 3). Hence, the association across domain was stronger in Study 2 than in Study 1, indicating reduced domain-specificity in expected benefit and risk perception. This finding suggests that the inter-correlations in risk perception and expected benefit ratings across domains are lower when domains are matched according to their range of values (i.e., low, medium, & high levels) on the rating scale. The Cronbach α scores—assessing internal consistency of the scales—were lower in Study 2 (Table 2) compared to Study 1 (Table 1), which may reflect greater diversity in the scale items due to categorizing items according to low, medium, high levels of expected benefits and risk perception in our scale construction.

Gender differences in expected benefits and risk-perception

We assessed gender differences in expected benefits and risk-perceptions with two domain (recreational, social, financial, health) × gender analyses of variance. Regarding expected benefits, there was no significant effect of gender ($M_{\text{women}} = 4.53$; $M_{\text{men}} = 4.68$; $F(1,254) = 0.80$, $p = .371$), but a significant effect of domain ($F(3,762) = 48.46$, $p < .001$, $\eta^2 = .16$), whereby benefits were expected to be greatest in the social domain ($M = 5.07$), followed by the health ($M = 4.71$), recreational ($M = 4.45$), and financial ($M = 4.20$) domains. There was no significant interaction between gender and domain ($F(3,762) = 1.73$, $p = .160$). Regarding risk perception, participants perceived the greatest risk in the financial domain ($M = 4.59$), followed by the recreational ($M = 4.02$), heath (M = 3.96), and social (M = 3.10) domains, qualified by a significant effect of domain ($F(3,762) = 109.09$, $p < .001$, $\eta^2 = .30$). While there was no significant effect of gender ($F(1,254) = 2.77$, $p = .098$), gender interacted with domain ($F(3,762) = 3.53$, $p = .019$, $\eta^2 = .01$). Follow-up independent samples t-tests confirmed that women perceived greater risk than men in the recreational ($M_{\text{women}} = 4.20$; $M_{\text{men}} = 3.84$; $t(254) = 2.00$, $p = .047$) and social ($M_{\text{women}} = 3.39$; $M_{\text{men}} = 2.82$; $t(254) = 2.51$, p

= .013) domains, but not in the financial (M_{women} = 4.68; M_{men} = 4.50; $t(254) = 0.74$, p = .461) and health (M_{women} = 4.00; M_{men} = 3.91; $t(254) = 0.55$, p = .585) domains.

Effects of expected benefits and risk perception on likelihood of engagement

To investigate the effects of expected benefits and risk perception on participants' self-reported likelihood of engagement, we conducted a multilevel linear regression analysis on participants' raw ratings of their likelihood of engagement. Fixed effects were included for expected benefit and risk perception ratings (as continuous predictors) and for risk domain and participant gender (as factors). A major advantage of the multilevel modeling approach over fixed effects regression modeling is that it enables the specification of random intercepts and random slopes. Allowing the intercepts to vary across participants (i.e., random intercepts) models individual differences in risk-taking, rather than assume that all participants have the same likelihood of engaging in the activities. Including random intercepts for risk domain models domain-specificity in risk-taking. Including random slopes for expected benefit and risk perception ratings models individual variability in their effects on likelihood of engagement. A further advantage of the multilevel modeling approach is that when random intercepts and random slopes are included, the correlation between the random effects can also be investigated. This enables us to assess, for example, whether greater sensitivity to expected benefit in determining one's likelihood of engagement is associated with greater (or smaller) sensitivity to one's risk perception. Hence, this approach circumvents a need to run multiple separate regression models on each individual participant for each risk domain.

Following the procedure proposed by Hoffman and Rovine (2007), we tested whether our model fit was significantly improved with the addition of random intercepts and random slopes by adding random effects to our model one-by-one, using the chi-square statistic to test for significant improvements in the model -2 log likelihood value. Accordingly, the model fit was significantly improved with the addition of random intercepts for participants

(χ^2 = 1202.90, p < .001) and additionally for risk domain (χ^2 = 27.81, p < .001). The model fit was further improved with the addition of random slopes for risk perception (χ^2 = 196.15, p < .001) and expected benefit (χ^2 = 308.95, p < .001).

The main effects model revealed that greater expected benefit (b = 0.38, t = 20.61, p < .001) and lower perceived risk (b = -0.36, t = 20.25, p < .001) were each associated with a higher self-reported likelihood of engaging in the activities (Table 4; Model 1). There was no significant effect of gender and the only effect of risk domain indicated a lower likelihood of engagement in the recreational domain (vs. the financial domain; b = -0.17, t = 1.99, p = .048; Table 4; Model 1). In Model 2 (Table 4), we included two-way interaction terms involving gender, risk perception, and risk domain. As predicted, risk perception interacted with gender (b = 0.12, t = 3.57, p < .001), indicating that women were more sensitive to their risk perceptions (i.e., less tolerant of perceived risk) than men. We also predicted that risk attitude would differ across risk domains. Risk perception interacted with the recreational domain (vs. the financial domain; b = -0.24, t = 9.32, p < .001), indicating that participants in general were more sensitive to their risk perceptions (i.e., less tolerant of risk) in the recreational domain. In Model 3 (Table 4), we included a three-way interaction term between gender, risk perception and risk domain. This further revealed that the gender difference in risk attitude (i.e., sensitivity to risk perception) was weaker in the health domain (vs. the financial domain; b = -0.13, t = 2.40, p = .016).

Figure 3 provides the best fitting slopes for risk perception on self-reported likelihood of engagement for men and women in the four risk domains, estimated from Model 3 (Table 4). Inspecting it, higher risk perception was associated with a lower likelihood of engaging in activities in each risk domain, but especially so in the recreational domain, indicating greater sensitivity to (or less tolerance of) perceived recreational risk. Moreover, women were more sensitive than men to their risk perceptions, indicating a lower tolerance of risk, but less so in the health domain (Figure 3). Conducting our regression model separately for each risk domain confirmed that gender interacted with risk perception in the recreational (b = 0.13, t = 3.29, p = .001), social (b = 0.13, t = 2.56, p = .010), and financial (b = 0.16, t = 3.48, p = .001) domains, but not in the health domain (b = 0.02, t = 0.33, p = .742).

Simple slope analysis, estimated from Model 3, confirmed that, when controlling for expected benefits, men reported a significantly higher likelihood than women of engaging in activities at 1*SD* above the overall mean risk perception rating in the recreational (b = 0.59, t = 2.58, p = .010), social (b = 0.64, t = 2.40, p = .017), and financial (b = 0.61, t = 2.67, p = .008) domains, but not in the health domain (b = 0.23, t = 0.91, p = .365). While Figure 3 shows that women reported that they were more likely than men to engage in low-risk recreational, social, and financial activities, our simple slope analysis indicated that women did not report a significantly higher likelihood than men of engaging in activities at 1*SD* below the mean risk perception rating in the recreational (b = -0.28, t = 1.30, p = .196), social (b = -0.21, t = 0.90, p = .370), financial (b = -0.37, t = 1.54, p = .125), or health (b = -0.01, t = 0.05, p = .959) domains.

The correlations between the random effects in our model reveals further insights into individual differences in risk-taking (Table 4; Model 3). First, the high negative associations between random intercepts and random slopes for risk perception and expected benefit indicate that participants who reported a higher overall likelihood of engaging in the activities were more sensitive to their perception of risk (i.e., a higher negative coefficient) and less sensitive to their expectations of benefit. This tendency may partly reflect a truncating effect of the upper and lower boundaries of the likelihood rating scale. Second, the random slopes for expected benefit and risk perception exhibited a strong positive correlation (Table 4; Model 3), indicating that greater sensitivity to the expected benefit of engaging in an activity was associated with less sensitivity (i.e., a weaker negative coefficient) to the perceived risk.

As such, people who were more attracted to the potential benefit of an activity were more tolerant of its potential risks. This latter finding replicates observations in other studies in the literature showing that the mean coefficient for risk perception correlates positively with the mean coefficient for expected benefit when regression models are instead conducted separately for each participant in a two-stage process (Weber et al., 2002).

In sum, the main findings of our multilevel regression modeling revealed that while participants were attracted to the benefits they expected of activities and were repelled by the risks they perceived, they were particularly repelled by recreational risks. Moreover, women were more repelled by risk than men, but less so in the health domain. Finally, greater positive sensitivity to the expected benefit of activities (i.e., finding the benefit more alluring) was associated with reduced negative sensitivity to the perceived risk (i.e., greater tolerance of risk).

Study 3

In Study 3, we employed a large sample of participants to revise the scale items of our new risk-taking scale to improve the scale reliability and structure. We also aimed to replicate our findings of Study 2 regarding gender and domain differences in perceived risk attitude.

Method

Participants

Five hundred sixty-eight participants residing in the US were recruited online using Amazon's Mechanical Turk via TurkPrime (Litman, Robinson, & Abberbock, 2016). Nine participants were excluded as they were located outside of the US, determined by their computer IP address. Our final sample included 559 participants (47% female, mean age = 38.01, *SD* = 12.00). Regarding education, 61 (11%) indicated high school as their highest educational attainment, 153 (27%) indicated some college, 255 (46%) indicated university, and 90

(16%) indicated postgraduate education. Regarding employment, the majority were in full- or part-time employment (n = 450; 81%).

Materials and procedure

Participants completed the 36-item scale developed in Study 2. As the item-total correlations in Study 2 revealed some items were poorly correlated with their within-domain subscale total scores, we also included four additional items in the recreational domain, one additional item in the financial domain, and one additional item in the health domain. The additional items were selected that matched the expected benefit and perceived risk category (i.e., low, medium, high) of the problematic item in each domain. As in Study 2, we selected items that were generated most frequently by participants in Study 1. Following the procedure introduced in Study 2, participants were randomly assigned to first complete either the risk perception or expected benefits section before completing the likelihood rating section.

After rating the 42 scale items, participants were shown the scale items in a list in a random order and were asked to group them according to their domain of life. Participants were told that each item described an activity or behavior in some domain of life (e.g., health, social, recreational, financial etc.) and were asked to group the scale items according to whether they believed the items belonged to the same domain of life. To assign items to a domain, participants clicked and dragged related items with their mouse cursor to create a group. Participants were asked to assign all the items to a group. They could create up to 15 groups of related items, but were asked to create only as many groups as they needed to assign all the items. Each group could contain from one to as many items as they wished.

Results

Scale reliability

The six new items included as candidate replacement items exhibited higher item-total correlations than the problematic items identified in Study 2. Thus, we replaced the problematic items with the new items. Table 5 provides the item-total correlations within each domain for the three subscales of the final 36-item scale. The mean item-total correlations were generally improved in comparison with Study 2. In the social domain, ‘Engaging in casual sex’ and ‘Visiting family’ exhibited item-total correlations for one of the subscales that were slightly below .30. In the financial domain, ‘Opening a pension fund’ exhibited an item-total correlation below .30. In the health domain, ‘Healthy eating’ and ‘Engaging in unprotected sex’ both exhibited item-total correlations below .30 on two of the three subscales. Regarding Cronbach’s alpha, as shown in Table 2, the scores were generally improved across domains after replacing the problematic items, and were above .70 for the three subscales in each domain, except the health domain.

Scale structure

We employed an exploratory factor analysis (EFA) with oblique rotation (Direct Oblimin in SPSS Version 21) to investigate the structure of the final 36-item scale for the likelihood of engagement subscale. In our initial EFA, seven factors exhibited Eigenvalues that exceeded Kaiser’s criterion (i.e., >1) and together explained 55% of the variance. However, inspection of the scree plot of Eigenvalues showed an inflexion point that indicated a five-factor solution. Thus, in our final EFA we retained five factors, which together explained 51% of the variance.

Table 6 provides the pattern matrix of the rotated factor loadings. The structure matrix yielded a comparable pattern of rotated factor loadings. All recreational items loaded on Factor 1 (29% of the variance), indicated by factor loadings $\geq$.30 (highlighted in bold; Table 6), and two items loaded also on Factor 2 (9% of the variance). The social items showed a spread of loadings across factors. Five social items loaded on Factor 2, one item loaded on Factor 1, and four items loaded on Factor 4 (3% of the variance). The six financial items related to gambling loaded on Factor 3 (6% of the variance) and the three items related to investment loaded on Factor 2. One item loaded also on Factor 4. The health items showed a spread of loadings across multiple factors. Four health items loaded on Factor 2, one item loaded on Factor 1, two items loaded on Factor 4, and two items loaded on Factor 5 (3% of the variance). In sum, the recreational items loaded principally on one factor, the factor structure distinguished financial items related to gambling and investment, and the social and health items showed a spread of loadings across multiple factors.

To further explore the scale structure, we examined participants' judgments about the domains to which the items belong. Participants were asked to group related items that they believe belong to the same domain. They were not provided group labels (e.g., recreational domain) and could create up to ten groups to categorize all 36 items. As such, participants' judgments provide a comparison to our EFA. Four hundred seventy (of 559; 84%) participants assigned all 36 items to groups. We conducted our analysis on the data of participants who assigned all 36 items to groups. To explore participants' groupings, we conducted an agglomerative hierarchical clustering analysis using the "cluster" package in R (Maechler, Rousseeuw, Struyf, Hubert, & Hornik, 2018). In agglomerative hierarchical clustering, each scale item is first assigned to its own individual cluster. Based on a dissimilarity matrix of the pairwise dissimilarities (i.e., distances) between the clusters, the individual clusters are merged into increasingly inclusive clusters in a sequential process until all clusters have been merged into a single overarching cluster. At each stage, clusters separated by the shortest distance (i.e., are least dissimilar) in the dissimilarity matrix are combined. We employed complete-linkage clustering based on Gower distances in the dissimilarity matrix, which is suited to categorical data (Maechler et al., 2018). To determine the optimal number of clusters, we first inspected the Elbow plot of the sum of the squared distances within clusters, which provides a measure of the similarity of items within clusters (i.e., coherence within clusters).

This method indicated a five-cluster solution as increasing the number of clusters from five (*SS* = 3.59) to six (*SS* = 3.00) clusters led to relatively smaller reductions in the sum of squared within-cluster distances, compared to four clusters (*SS* = 4.86; Appendix C). Further inspection of a plot of the Silhouette coefficient, which provides a measure of within-cluster consistency, confirmed that within-cluster consistency was maximized by a five-cluster solution (coefficient = 0.50). Four-cluster (coefficient = 0.43) and six-cluster (coefficient = 0.49) solutions exhibited smaller Silhouette coefficients (Appendix C).

Figure 4 provides a dendrogram of the five clusters determined by our clustering analysis. The dendrogram height indicates the distance between clusters in the dissimilarity matrix, such that clusters with a shorter height are less dissimilar as they were more frequently grouped together in participants' judgments. The dendrogram also shows the hierarchical nature of our approach, whereby similar clusters are sequentially combined into increasingly inclusive clusters. As shown in Figure 4, all recreational items were contained in Cluster 3. Cluster 3 appears to reflect Factor 1 in our EFA, representing the recreational domain. Eight of the social items were contained in Cluster 4, which partially corresponds with Factor 2 in our EFA, representing the social domain. One item ('Engaging in casual sex') was assigned to Cluster 5, which loaded on Factor 4 in our EFA, and exhibited below mean item-total correlations for the three subscales. Hence, this item may be the least relevant to the Social domain. The six financial items related to gambling were contained in Cluster 1 and the three items related to investment were contained in Cluster 2. This distinction maps nicely onto our EFA, in which the gambling items loaded on Factor 3 and the investment items loaded on Factor 2. The distinction between financial gambling and investment items also resonates with a similar distinction made by Weber et al. (2002) in their development of the DOSPERT scale. In our dendrogram, Clusters 1 and 2 were combined in a superordinate cluster, which appears to represent the broader financial domain. Thus, the financial domain can be considered to contain financial investment and gambling sub-domains. Six of the health items were contained in Cluster 5, indicating a health domain. The three remaining health items were contained in Cluster 3 and thus may be more related to behavior in the recreational domain. This finding resonates with our EFA, in which ‘Running’ also loaded on Factor 1 with other recreational items.

In sum, our EFA and clustering analysis provide support for the scale structure, but also indicate possible refinements. The recreational items loaded principally on a single factor in our EFA and were all contained in a single cluster in our clustering analysis, indicating a coherent recreational domain. The social items showed a spread of loadings across factors in our EFA, but were more coherent in our clustering analysis. The Financial domain appears to comprise financial gambling and investment sub-domains. Weber et al. (2002) came to the same conclusion with a very different approach to scale construction. They constructed their scale from existing risk-taking measures, whereas we asked participants to generate their own examples of activities and behaviors. Our findings provide convergent support for these distinct sub-domains within financial risk-taking. Finally, the health items loaded on various factors in our EFA and were spread across clusters in our clustering analysis, which indicates that the health items do not provide a coherent set of items. The overlap between the health domain and recreational domain in our clustering analysis and EFA suggests that some health items, namely ‘Running’ and ‘Exercising at the gym’ have both a recreational and health component. We probed whether the reliability of items in the health domain could be improved by using the scale structure revealed in our clustering analysis as determined by participants’ judgments about the domains to which the items belong. However, scale reliability was not improved for the expected benefit (Cronbach’s alpha = 0.73; item-total correlation = 0.49, range = -0.16-0.68), risk perception (Cronbach’s alpha = 0.69; item-total correlation =

0.44, range = 0.27-0.60), or likelihood of engagement (Cronbach's alpha = 0.58; item-total correlation = 0.32, range = -0.07-0.46) subscales.

Gender differences in expected benefits and risk perception

As in Study 2, we assessed gender differences in expected benefits and risk-perceptions with two domain (recreational, social, financial, health) × gender analyses of variance. Regarding expected benefits, men (M = 4.88) expected slightly greater benefits to taking a risk compared to women (M = 4.60; $F(1,557) = 5.73$, $p = .017$, $\eta^2 = .01$). An effect of domain indicated differences in expected benefits ($F(3,1671) = 39.70$, $p < .001$, $\eta^2 = .07$), which were perceived to be greatest in the social domain (M = 5.02), followed by the recreational (M = 4.78), health (M = 4.59), and financial (M = 4.58) domains. There was no significant interaction between gender and domain ($F(3,1671) = 1.37$, $p = .251$). Regarding risk perception, participants perceived the greatest risk in the financial domain (M = 4.48), followed by the health (M = 4.27), recreational (M = 4.08), and social (M = 3.52) domains, qualified by a significant effect of domain ($F(3,1671) = 134.90$, $p < .001$, $\eta^2 = .20$). There was no significant effect of gender ($F(1,557) = 0.41$, $p = .524$), and no interaction with domain ($F(3,1671) = 2.24$, $p = .081$).

Effects of expected benefits and risk perception on likelihood of engagement

We employed the multilevel linear regression analysis introduced in Study 2 to investigate effects of expected benefits and risk perception on self-reported likelihood of engagement. As in Study 2, fixed effects were included for expected benefits and risk perceptions ratings (as continuous predictors) and for risk domain and participant gender (as factors). Random intercepts were included for participants and risk domain, and random slopes were included for risk perception and expected benefits. The model fit was significantly improved with the addition of random intercepts for participants ($\chi^2 = 2869.05$, $p < .001$) and additionally for risk domain ($\chi^2 = 313.31$, $p < .001$). The model fit was further improved with the addition of random slopes for risk perception ($\chi^2 = 464.00$, $p < .001$) and expected benefits ($\chi^2 = 650.56$, $p < .001$).

In our main effects model (Table 7; Model 1), greater expected benefits ($b = 0.39$, $t = 31.08$, $p < .001$) and lower perceived risk ($b = -0.23$, $t = 19.47$, $p < .001$) were each associated with a higher self-reported likelihood of engaging in the activities. There was a significant main effect of gender ($b = 0.32$, $t = 3.02$, $p = .003$), such that men ($M = 4.94$) reported a higher likelihood than women ($M = 4.53$) of engaging in the activities. There were also significant effects of risk domain, whereby participants indicated a higher likelihood of engaging in health ($b = 0.47$, $t = 7.45$, $p < .001$) and social ($b = 0.15$, $t = 2.27$, $p = .024$) activities and a lower likelihood of engaging in recreational activities ($b = -0.71$, $t = 11.03$, $p < .001$) in comparison with financial activities. Model 2 (Table 7) included two-way interaction terms involving gender, risk perception, and risk domain. Importantly, this replicated our finding from Study 2 regarding an interaction between risk perception and gender ($b = 0.11$, $t = 4.94$, $p < .001$), indicating that women were more sensitive to their risk perceptions (i.e., less tolerant of perceived risk) than men. Further replicating our findings of Study 2, risk attitude differed across risk domains, as indicated by interactions between risk perception and the health ($b = -0.14$, $t = 7.43$, $p < .001$), recreational ($b = -0.14$, $t = 7.16$, $p < .001$), and social ($b = -0.19$, $t = 9.52$, $p < .001$) domains in comparison with the financial domain. Gender interacted with the recreational domain ($b = 0.44$, $t = 3.47$, $p = .001$), whereby men's tendency to report a higher likelihood than women of engaging in activities was greater in the recreational domain ($M_{\text{men}} = 4.32$; $M_{\text{women}} = 3.57$) compared to the financial domain ($M_{\text{men}} = 4.40$; $M_{\text{women}} = 4.03$). Model 3 (Table 7) included a three-way interaction term between gender, risk perception and risk domain. This revealed interactions in the recreational ($b = -0.08$, $t = 2.15$, $p = .031$) and health ($b = -0.11$, $t = 2.95$, $p = .003$) domains in comparison with the financial domain, indicating that the gender difference in risk attitude (i.e., sensitivity to risk perception) was weaker in the health and recreational domains (vs. the financial domain). Our scale reliability analysis indicated lower reliability in health domain. The findings of our multilevel linear regression analysis were not altered by removing the two items in the health domain that exhibited item-total correlations below 0.30.

Figure 5 provides the best fitting slopes for risk perception on self-reported likelihood of engagement, estimated from Model 3 (Table 7). Higher risk perception was associated with a lower reported likelihood of engaging in activities in each risk domain. Women were more sensitive than men to their risk perceptions, implying lower risk tolerance among women, but least so in the health domain (Figure 5). Conducting our regression model separately for each risk domain confirmed our findings of Study 2, such that gender interacted with risk perception in the recreational ($b = 0.07$, $t = 2.23$, $p = .026$), social ($b = 0.14$, $t = 4.34$, $p < .001$), and financial ($b = 0.16$, $t = 4.84$, $p < .001$) domains, but not in the health domain ($b = 0.06$, $t = 1.73$, $p = .085$). Simple slope analysis, estimated from Model 3, showed that, when controlling for expected benefits, men reported a significantly higher likelihood than women of engaging in activities at 1*SD* above the overall mean risk perception rating in the recreational ($b = 0.93$, $t = 5.71$, $p < .001$), social ($b = 0.85$, $t = 4.75$, $p < .001$), financial ($b = 0.71$, $t = 4.33$, $p < .001$), and health ($b = 0.39$, $t = 2.30$, $p = .022$) domains. While Figure 5 also shows that women reported a higher likelihood than men of engaging in low-risk social and financial activities, our simple slope analysis indicated that women did not report a significantly higher likelihood than men of engaging in activities at 1*SD* below the mean risk perception rating in the social ($b = -0.13$, $t = 0.90$, $p = .370$) and financial ($b = -0.25$, $t = 1.47$, $p = .143$) domains.

General Discussion

Within the psychological risk-return framework, researchers have concluded that risk attitude is stable across gender and risk domains (e.g., Weber et al., 2002; Hanoch et al., 2006). We proposed that the DOSPERT—the principle measure used to assess domain-specificity within this framework—may have failed to detect differences in risk attitude by failing to represent all combinations of levels of risk perception and expected benefit (i.e., [low, medium, high] risk perception; [low, medium, high] expected benefit). To remedy this problem, we had participants in Study 1 generate their own examples of activities and in Study 2 and 3 we selected among the generated activities to construct a set of items for each of four risk domains that captured various levels of expected benefit and risk perception.

Gender differences in risk perception, expected benefit, perceived risk-attitude

Previous research using the DOSPERT has suggested that differences in risk-taking behavior across gender and risk domain are due to differences in perceptions of risk and benefit, conditioned by culture (Weber & Johnson, 2008). Women are seemingly more cautious and perceive greater risk in most domains (e.g., recreational, health, financial), except in the social domain (Hanoch et al., 2006; Rolison et al., 2014; Weber et al., 2002). The original items of the DOSPERT scale were devised by researchers and drawn from various existing scales (Weber et al., 2002). However, Morgenroth, Fine, Ryan, and Genat (2018) have argued that some DOSPERT items are more characteristic of male than female behaviors and thus normatively bias the scale toward showing greater risk-taking tendencies among men. The authors found that gender differences in financial risk-taking disappeared when the original items were replaced with new gender-neutral or stereotypically female items (Morgenroth et al., 2018). In the recreational domain, gender differences similarly disappeared for the new items. In the health and social domains, women actually indicated higher likelihoods of risk-taking than men for the new items. Their findings indicate that apparent gender differences in risk-taking can be highly dependent on the specific items selected to represent a risk domain. Moreover, Zhang et al. (in press) found that the underlying factor structure of the DOSPERT differs for men and women, which raises additional concerns about the assessment of gender differences in self-reported risk-taking on the DOSPERT.

We took a very different approach to our scale construction by asking participants in Study 1 to generate their own examples of activities in each domain. In Study 2, we found that women perceived greater risk than men in the recreational and social domains, but not in the health and financial domains. In Study 3, our refined scale showed no gender differences in risk perception. In Studies 2 and 3, our multilevel modeling analysis revealed that risk attitude—measured as the coefficient for the effect of risk perception on self-reported likelihood of engagement, controlling for expected benefit—differed between men and women. Specifically, women were more sensitive to their risk perceptions (i.e., less tolerant of perceived risk) than men in all risk domains, except the health domain. This finding challenges the view that perceived risk attitude is stable across gender (e.g., Weber et al., 2002). Men reported a higher likelihood than women of engaging in activities that were perceived to be high-risk. Women did not report a lower likelihood than men of engaging in activities that were perceived to be low-risk. Therefore, apparent gender differences in risk-taking behavior (e.g., in a particular risk domain) may depend on how risky the items are, such that men can appear more risk-taking than women for high-risk items, whereas gender differences may disappear for low-risk items. As such, this tendency, driven by gender differences in risk attitude may partially explain apparent domain-specificity in gender differences in risk taking.

Domain-specificity in risk perception, expected benefit, perceived risk-attitude

Our results show for the first time, within the psychological risk-return framework, that attitudes toward perceived risk differ across risk domains when controlling for expected benefits. Moreover, gender differences in perceived risk attitude further depended on domain. In Studies 2 and 3, we found that gender differences in perceived risk attitude were weaker in the health domain than in other domains, indicating that as perceived risk increased women were not significantly less tolerant than men of risks they perceived. High risk items in the health domain included ‘Having an invasive health procedure (e.g., colonic irrigation)’, ‘Engaging in unprotected sex’, and ‘Consuming sugar’. It appears that women are no less tolerant than men of perceived risk to attain the expected benefits of engaging in such high-risk activities. As discussed later, however, domain differences involving the health domain should be considered with caution as the health domain exhibited lower scale reliability than other domains. We speculate that risk perception and expected benefits may be context-dependent. People may be more tolerant of risk to attain certain kinds of benefits. For instance, the benefits of health-related activities (e.g., ‘Exercising at the gym’) may not be equivalent to the benefits of engaging in activities in other domains (e.g., ‘Skiing’; recreational domain) even when they receive a similar rating on an expected benefits scale. Similarly, activities that receive equivalent ratings of perceived risk may differ in qualitative aspects of people’s risk perception. For example, activities may differ in the time horizon of their perceived negative outcomes: ‘Playing poker’ (financial domain) could result in an immediate financial loss, whereas the negative consequences of ‘Running’ (health domain) are less clear and may have a longer time horizon (e.g., future joint injuries). Such nuanced differences in expected benefits and perceived risks may partially underlie domain and gender differences in perceived risk attitude. A fruitful avenue for future research would be to explore the nuanced nature of expected benefits and risk perception, such as in terms of their anticipated time horizons, moving beyond singular rating scales.

Unlike previous studies (e.g., Weber et al., 2002; Rolison et al., 2014), we found only small domain differences in risk perception and expected benefit ratings. In fact, the inter-correlations in risk perception and expected benefit ratings across domains were higher in Study 2 than in Study 1 when we matched the domains according to low, medium, and high levels on the rating scale. This finding suggests that domain-specificity in perceptions of risks and benefits may have been exaggerated—in terms of dissociations across risk domains—in previous studies with the DOSPERT. Recently, researchers have questioned the extent of domain-specificity in risk-taking behavior (Highhouse et al., 2017; Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017; Zhang et al., in press). Frey et al. (2017) assessed individual differences in risk-taking across 39 measures of risk preference. Their analysis revealed evidence of a general trait in risk preference across measures in addition to domain-specific risk-taking tendencies. Highhouse et al. (2017) employed a form of confirmatory factor analysis known as bifactor analysis to assess the factor structure of the DOSPERT. Their bifactor analysis enabled them to assess simultaneously domain-general and domain-specific risk behavior and yielded evidence for both a domain-general factor and domain-specific factors underlying risk behavior on the scale. Thus, risk-taking behavior across domains may involve both domain-specific tendencies as well as general risk-taking tendencies. Our current findings speak to this ongoing debate. We found that domain differences in perceived risk attitude emerged when domains were matched according to their range of perceived risks and benefits. This finding implies domain-specificity in perceived risk attitude, such that people are more tolerant of risk to attain a benefit in some domains than in others. We expect that domain-specificity in perceived risk attitude partly underlies domain differences in risk preference observed in other studies (e.g., Highhouse et al., 2017), in addition to domain differences in perceived risks and benefits.

The structure of risk domains

In Study 3, we asked participants to group the items of our risk-taking scale according to the domains that they believed the items belong. Participants could create up to ten groups to categorize all the scale items and were not guided by domain labels (e.g., recreational domain). This approach enabled us to compare participants' judgments of the items with our exploratory factor analysis on their self-reported likelihood of engaging in the activities. Our clustering analysis of participants' judgments and our factor analysis both indicated that the financial domain comprises investment and gambling sub-domains. This distinction resonates with a similar distinction made by Weber et al. (2002) in their construction of the DOSPERT. Moreover, in their bifactor analysis of the DOSPERT, Highhouse et al. (2017) found that when controlling for financial risk-taking propensity in general, likelihood of engagement ratings for investment items were negatively correlated with ratings for gambling items. Thus, investment and gambling activities appear to belong to sub-domains of financial risk-taking. Our current findings show compelling support for this distinction based on scale items that participants generated themselves from their own experiences.

In our factor analysis, items in the health domain loaded on various factors. Our clustering analysis on participants' judgments further showed that some items participants had generated for the health domain, such as 'Running' and 'Exercising at the gym', were more often associated with recreational activities. Compared to other domains, the health items also showed the lowest Cronbach alpha scores for internal consistency and the lowest item-total correlations for the three sub-scales. Together, these findings suggest that the health domain is the least coherent domain as some health-related activities have features that are relevant also to other domains, such as the recreational domain. In their factor analysis of the DOSPERT, Weber et al. (2002) also observed that some of the DOSPERT health items loaded on multiple factors.

The negative association between risk perception and expected benefit

In Study 1, when participants were asked to generate their own examples of activities, we found that risk perception was negatively correlated with expected benefits. This tendency for an inverse relationship between risk and benefit judgments is well-documented in the literature (e.g., Alhakami & Slovic, 1994; Slovic, 1997; Weber et a., 2002). One proposed explanation is that people judge the risks and benefits of an activity (or decision option) on their affective evaluation of the activity (Alhakami & Slovic, 1994; Finucane, Alhakami, Slovic, & Johnson, 2000; Slovic & Peters, 2006). This research suggests that when people experience positive feelings toward an activity, they perceive the risks to be low and the benefits to be high, and when they experience negative feelings toward an activity they perceive the risks to be high and the benefits to be low. This inverse relationship contrasts with a positive association between risks and benefits in some situations—e.g., financial contexts where high-risk investments often yield higher potential returns than low-risk investments. A serendipitous finding of our current research is that when participants were asked to freely generate examples of activities (i.e., control condition, Study 1) they tended to generate activities that they perceived to be very low in risk (except for the financial domain). This finding highlights that people may often face situations in which they perceive that risk has little bearing on their decisions.

Conclusion

In conclusion, in contrast with previous studies, our findings reveal gender and risk domain differences in attitudes toward perceived risk. Namely, women were more sensitive to their risk perceptions in the recreational, social, and financial domains and thus less tolerant of risk than men. Domain differences in risk attitude revealed that people are more sensitive to risk in some domains than in others. These findings help bridge a gap between the psychological risk-return framework and models of choice behavior within the expected utility framework where men have often been found to exhibit higher levels of risk tolerance than women in their risky choice behavior.

References

Alhakami, A. S., & Slovic, P. (1994). A psychological study of the inverse relationship between perceived risk and perceived benefit. *Risk analysis, 14,* 1085–1096.

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

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computer and Human Behavior, 29,* 2156–2160.

Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior and Organization, 83,* 50–58.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science, 20,* 211–216.

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

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances, 3,* e1701381.

Gibson, E., Piantadosi, S., & Fedorenko, K. (2011). Using mechanical turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass, 5*, 509–524.

Hanoch, Y., Johnson, J. G., & Wilke, A. (2006). Domain specificity in experimental measures and participant recruitment an application to risk-taking behavior. *Psychological Science, 17,* 300–304.

Harris, C. R., & Jenkins, M., & Glaser, D. (2006). Gender differences in risk assessment: Why do women take fewer risks than men? *Judgment and Decision Making, 1,* 48–63.

Highhouse, S., Nye, C. D., Zhang, D. C., & Rada, T. B. (2017). Structure of the Dospert: Is there evidence for a general risk factor? *Journal of Behavioral Decision Making, 30,* 400–406.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods, 39,* 101–117.

Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Levin, I. P., Snyder, M. A., & Chapman, D. P. (1988). The interaction of experiential and situational factors and gender in a simulated risky decision-making task. *The Journal of Psychology, 122,* 173–181.

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: a versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods,* 1 - 10.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2018). *Cluster: cluster analysis basics and extensions*. R package version 2.0.7-1

Morgenroth, T., Fine, C., Ryan, M. K., & Genat, A. E. (2018). Sex, drugs, and reckless driving: Are measures biased toward identifying risk-taking in men? *Social Psychological and Personality Science, 9,* 744–753.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed). New York: McGraw-Hill.

Powell, M., & Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology, 18,* 605–628.

Preacher, K. J. (2002, May). Calculation for the test of the difference between two independent correlation coefficients [Computer software]. Available from http://quantpsy.org.

Rolison, J. J., Hanoch, Y., & Freund, A. M. (in press). Perception of risk for older adults: Differences in evaluations for self versus others and across risk domains. *Gerontology,* 1-13.

Rolison, J. J., Hanoch, Y., Wood, S., & Pi-Ju, L. (2014). Risk taking differences across the adult lifespan: A question of age and domain. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 69,* 870–880.

Rosen, A. B., Tsai, J. S., & Downs, S. M. (2003). Variations in risk attitude across race, gender, and education. *Medical Decision Making, 23,* 511–517.

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences, 106,* 15268–15273.

Slovic, P. (1997). Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield. In *Psychological Perspectives to Environmental and Ethical Issues in Management* (pp. 277–313), Bazerman M, Messick D, Tenbrunsel A, Wade-Benzoni K (eds.). Jossey-Bass: San Francisco, CA.

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

Weber, E. U. (1997). The utility of measuring and modeling perceived risk. In *Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce* (pp. 45–57), Marley AAJ (ed.). Erlbaum: Mahwah, NJ.

Weber, E. U. (1998). Who's afraid of a little risk? New evidence for general risk aversion. In *Decision Research from Bayesian Approaches to Normative Systems: Reflections on the Contributions of Ward Edwards*, Shanteau J, Mellers BA, Schum D (eds.). Kluwer Academic Press: Norwell, MA.

Weber, E. U., Blais, A-R, & Betz, N. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15,* 263–290.

Weber, E. U., & Hsee, C. (1998). Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk. *Management Science, 44,* 1205–1217.

Weber, E. U., & Hsee, C. K. (1999). Models and mosaics: Investigating cross-cultural differences in risk perception and risk preference. *Psychonomic Bulletin and Review, 6,* 611–617.

Weber, E. U., & Milliman, R. A. (1997). Perceived risk attitudes: Relating risk perception to risky choice. *Management science, 43,* 123–144.

Zhang, D. C., Foster, G. C., & McKenna, M. G. (in press). Is the DOSPERT gender invariant? A psychometric test of measurement invariance. *Journal of Behavioral Decision Making*.

Table 1. Study 1: Inter-correlations across domains for expected benefits and risk perceptions

	Expected benefit			
	Recreational	Social	Financial	Health
Recreational	(.92)			
Social	.60***	(.91)		
Financial	.21*	.42***	(.88)	
Health	.55***	.67***	.31**	(.96)

	Risk perception			
	Recreational	Social	Financial	Health
Recreational	(.94)			
Social	.56***	(.92)		
Financial	.04	.18	(.85)	
Health	.73***	.60***	.21*	(.94)

Note. *$p \leq .05$, **$p \leq .01$, ***$p \leq .001$, 2-tailed significance test of the Pearson *r* correlation coefficient compared to zero. Cronbach α values are in parenthesis.

Table 2: Studies 2 and 3: Cronbach's alphas and mean within-domain item-total correlations for expected benefit, risk perception, and likelihood of engagement subscales.

	Cronbach's alpha			Item-total correlation		
Domain	Expected benefits	Risk perception	Likelihood of engagement	Expected benefits	Risk perception	Likelihood of engagement
Study 2						
Recreational	0.80	0.76	0.72	0.52 (0.29-0.69)	0.47 (0.23-0.61)	0.42 (0.10-0.65)
Social	0.79	0.89	0.81	0.52 (0.32-0.64)	0.67 (0.46-0.77)	0.53 (0.41-0.69)
Financial	0.79	0.88	0.80	0.52 (-0.01-0.70)	0.65 (0.18-0.80	0.52 (0.10-0.45)
Health	0.63	0.72	0.62	0.34 (0.18-0.47)	0.43 (0.16-0.58)	0.33 (0.12-0.45)
Study 3						
Recreational	0.88	0.83	0.87	0.64 (0.50-0.74)	0.57 (0.32-0.69)	0.62 (0.32-0.69
Social	0.83	0.88	0.78	0.56 (0.37-0.67)	0.65 (0.25-0.76)	0.50 (0.26-0.63)
Financial	0.88	0.87	0.86	0.64 (0.23-0.81)	0.63 (0.37-0.75)	0.61 (0.47-0.74)
Health	0.69	0.79	0.61	0.39 (0.19-0.49)	0.51 (0.03-0.67)	0.32 (0.19-0.4)

Note. Values in parenthesis indicate the minimum and maximum values

Table 3. Study 2: Inter-correlations across domains for expected benefits, risk perception, and likelihood of engagement

	Expected benefits			
	Recreational	Social	Financial	Health
Recreational	(.82)			
Social	.77***†††	(.82)		
Financial	.63***†††	.62***††	(.82)	
Health	.70***††	.71***	.61***†††	(.66)
	Risk perception			
	Recreational	Social	Financial	Health
Recreational	(.78)			
Social	.80***†††	(.90)		
Financial	.64***†††	.58***†††	(.89)	
Health	.80***	.82***†††	.68***†††	(.75)
	Likelihood of engagement			
	Recreational	Social	Financial	Health
Recreational	(.75)			
Social	.70***	(.83)		
Financial	.59***	.56***	(.82)	
Health	.72***	.75***	.58***	(.65)

Note. *$p \leq .05$, **$p \leq .01$, ***$p \leq .001$, 2-tailed significance test of the Pearson r correlation coefficient compared to zero; $^{\dagger}p \leq .05$, $^{\dagger\dagger}p \leq .01$, $^{\dagger\dagger\dagger}p \leq .001$, 2-tailed significance test of the Pearson r correlation coefficient compared to the corresponding correlation coefficient in Study 1; Cronbach α values are in parenthesis.

Table 4. Study 2: Multilevel linear regression analysis on likelihood of engagement

	Model 1	Model 2	Model 3
Intercept	4.26***	4.07***	4.19***
Male gender	0.09	-0.30	-0.51
Expected benefits	0.38***	0.39***	0.39***
Risk perception	-0.36***	-0.35***	-0.38***
Health domain	0.13	0.15	-0.16
Recreational domain	-0.17*	0.77***	0.71**
Social domain	-0.10	0.04	-0.05
Male gender × risk perception		0.12***	0.16***
Male gender × health domain		-0.05	0.46
Male gender × recreational domain		0.01	0.11
Male gender × social domain		0.05	0.20
Risk perception × health domain		0.02	0.10*
Risk perception × recreational domain		-0.24***	-0.23***
Risk perception × social domain		-0.03	-0.01
Male gender × risk perception × health domain			-0.13*
Male gender × risk perception × recreational domain			-0.02
Male gender × risk perception × social domain			-0.03
Goodness of fit			
-2 log likelihood	42987.42	42844.60	42837.12
-2 log likelihood change†		142.82***	7.47

Random coefficients (Model 3)

	SD	Corr (intercepts)	Corr (risk perception)
Intercepts	1.60		
Risk perception	0.13	-.94	
Expected benefits	0.22	-.96	.97

Note. $^*p \leq .05$, $^{**}p \leq .01$, $^{***} p \leq .001$; †Change in relation to previous model.

Table 5: Study 3: Final Questionnaire items

				Item-total correlations		
Risk Domain	Questionnaire Item	Benefit level	Risk level	Expected benefits	Risk perception	Likelihood of engagement
Recreational	1. Going white-water rafting	High	High	0.74	0.51	0.73
	2. Hiking	High	Medium	0.53	0.66	0.49
	3. Cycling	High	Low	0.5	0.69	0.55
	4. Going skydiving	Medium	High	0.64	0.32	0.65
	5. Skiing	Medium	Medium	0.71	0.63	0.7
	6. Fishing	Medium	Low	0.56	0.63	0.54
	7. Rock climbing	Low	High	0.72	0.52	0.72
	8. Going scuba diving	Low	Medium	0.7	0.6	0.71
	9. Bowling	Low	Low	0.67	0.59	0.52
Social	1. Meeting new people	High	High	0.63	0.74	0.57
	2. Socializing with friends	High	Medium	0.55	0.73	0.48
	3. Visiting family	High	Low	0.37	0.69	0.26
	4. Partying	Medium	High	0.62	0.61	0.63
	5. Dating	Medium	Medium	0.58	0.71	0.59
	6. Joining a social club	Medium	Low	0.67	0.76	0.57
	7. Engaging in casual sex	Low	High	0.43	0.25	0.36
	8. Going to a bar	Low	Medium	0.55	0.66	0.55
	9. Chatting online	Low	Low	0.65	0.71	0.46
Financial	1. Investing in the stock market	High	High	0.50	0.53	0.50
	2. Making a financial investment	High	Medium	0.41	0.48	0.47
	3. Opening a pension fund	High	Low	0.23	0.37	0.47
	4. Gambling at a casino	Medium	High	0.76	0.66	0.74
	5. Playing poker	Medium	Medium	0.74	0.68	0.62
	6. Playing a scratch card	Medium	Low	0.79	0.75	0.64
	7. Playing a slot machine	Low	High	0.81	0.75	0.74
	8. Playing the lottery	Low	Medium	0.76	0.75	0.68
	9. Gambling on a sporting event	Low	Low	0.76	0.7	0.63
Health	1. Having an invasive health procedure (e.g., colonic irrigation)	High	High	0.41	0.45	0.35
	2. Exercising at the gym	High	Medium	0.34	0.65	0.4

3. Healthy eating	High	Low	0.19	0.65	0.25
4. Engaging in unprotected sex	Medium	High	0.41	0.03	0.19
5. Running	Medium	Medium	0.42	0.63	0.39
6. Walking	Medium	Low	0.33	0.67	0.3
7. Consuming sugar	Low	High	0.48	0.58	0.31
8. Drinking alcohol	Low	Medium	0.49	0.48	0.37
9. Unhealthy eating (e.g., eating a hamburger)	Low	Low	0.42	0.45	0.31

Table 6: Study 3: Factor loadings of the 36-item for the likelihood of engagement subscale

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Going white-water rafting (R)	**0.81**	-0.02	0.08	-0.09	0.06
Hiking (R)	**0.37**	**0.46**	-0.01	-0.08	0.06
Going skydiving (R)	**0.76**	-0.19	0.04	0.10	-0.01
Skiing (R)	**0.68**	0.02	-0.01	0.15	-0.03
Going scuba diving (R)	**0.79**	-0.02	-0.01	0.02	0.06
Cycling (R)	**0.33**	**0.41**	0.01	0.16	-0.08
Fishing (R)	**0.40**	0.20	0.26	-0.10	0.00
Rock climbing (R)	**0.69**	-0.01	0.02	0.17	-0.09
Bowling (R)	**0.36**	0.23	0.26	-0.08	0.08
Meeting new people (S)	0.11	**0.63**	0.00	0.14	0.11
Socializing with friends (S)	-0.06	**0.67**	-0.05	0.13	0.21
Visiting family (S)	-0.09	**0.61**	0.03	-0.04	0.07
Partying (S)	0.04	0.09	0.21	**0.65**	-0.01
Dating (S)	0.08	0.21	-0.02	**0.55**	-0.02
Joining a social club (S)	**0.31**	**0.35**	0.07	0.26	-0.18
Engaging in casual sex (S)	0.22	-0.24	-0.06	**0.62**	0.05
Going to a bar (S)	-0.07	0.14	0.16	**0.66**	0.05
Chatting online (S)	0.08	**0.34**	0.00	0.21	0.19
Investing in the stock market (F)	0.09	**0.41**	0.20	0.18	-0.17
Making a financial investment (F)	0.11	**0.52**	0.17	0.07	-0.13
Gambling at a casino (F)	-0.09	-0.09	**0.81**	0.22	-0.01
Playing poker (F)	0.19	0.01	**0.43**	0.28	-0.07
Playing a scratch card (F)	0.04	-0.02	**0.79**	-0.09	0.14
Playing a slot machine (F)	-0.07	0.00	**0.86**	0.10	-0.02
Playing the lottery (F)	0.03	0.01	**0.81**	-0.12	0.13
Gambling on a sporting event (F)	0.20	-0.12	**0.45**	**0.35**	-0.14
Opening a pension fund (F)	0.24	**0.41**	0.21	-0.02	-0.09
Having an invasive health-promoting procedure (e.g., colonic irrigation) (H)	0.23	0.06	0.17	0.28	-0.02
Exercising at the gym (H)	0.18	**0.52**	0.03	0.11	-0.16
Healthy eating (H)	-0.07	**0.73**	0.00	-0.09	-0.10
Engaging in unprotected sex (H)	0.24	-0.29	0.06	**0.40**	0.19
Running (H)	**0.32**	**0.37**	-0.04	0.20	-0.12
Walking (H)	-0.16	**0.74**	-0.06	-0.11	0.17
Drinking alcohol (H)	-0.11	0.06	0.19	**0.50**	0.18
Unhealthy eating (e.g., eating a hamburger) (H)	0.08	-0.02	0.09	0.07	**0.72**
Consuming sugar (H)	0.01	0.13	0.07	0.06	**0.58**

Table 7. Study 3: Multilevel linear regression analysis on likelihood of engagement

	Model 1	Model 2	Model 3
Intercept	3.48***	3.22***	3.35***
Male gender	0.32**	-0.21	-0.45*
Expected benefits	0.39***	0.39***	0.39***
Risk perception	-0.23***	-0.17***	-0.19***
Health domain	0.47***	1.12***	0.89***
Recreational domain	-0.71***	-0.30*	-0.48**
Social domain	0.15*	0.92***	0.89***
Male gender × risk perception		0.11***	0.16***
Male gender × health domain		0.01	0.48*
Male gender × recreational domain		0.44***	0.82***
Male gender × social domain		0.06	0.11
Risk perception × health domain		-0.14***	-0.09**
Risk perception × recreational domain		-0.14***	-0.10***
Risk perception × social domain		-0.19***	-0.19***
Male gender × risk perception × health domain			-0.11**
Male gender × risk perception × recreational domain			-0.08*
Male gender × risk perception × social domain			0.00
Goodness of fit			
-2 log likelihood	92313.50	92175.2	92160.32
-2 log likelihood change†		138.30***	14.90**

Random coefficients (Model 3)

	SD	Corr (intercepts)	Corr (risk perception)
Intercepts	1.67		
Risk perception	0.13	-.70	
Expected benefits	0.21	-.90	.59

Note. *$p \leq .05$, **$p \leq .01$, *** $p \leq .001$; †Change in relation to previous model.

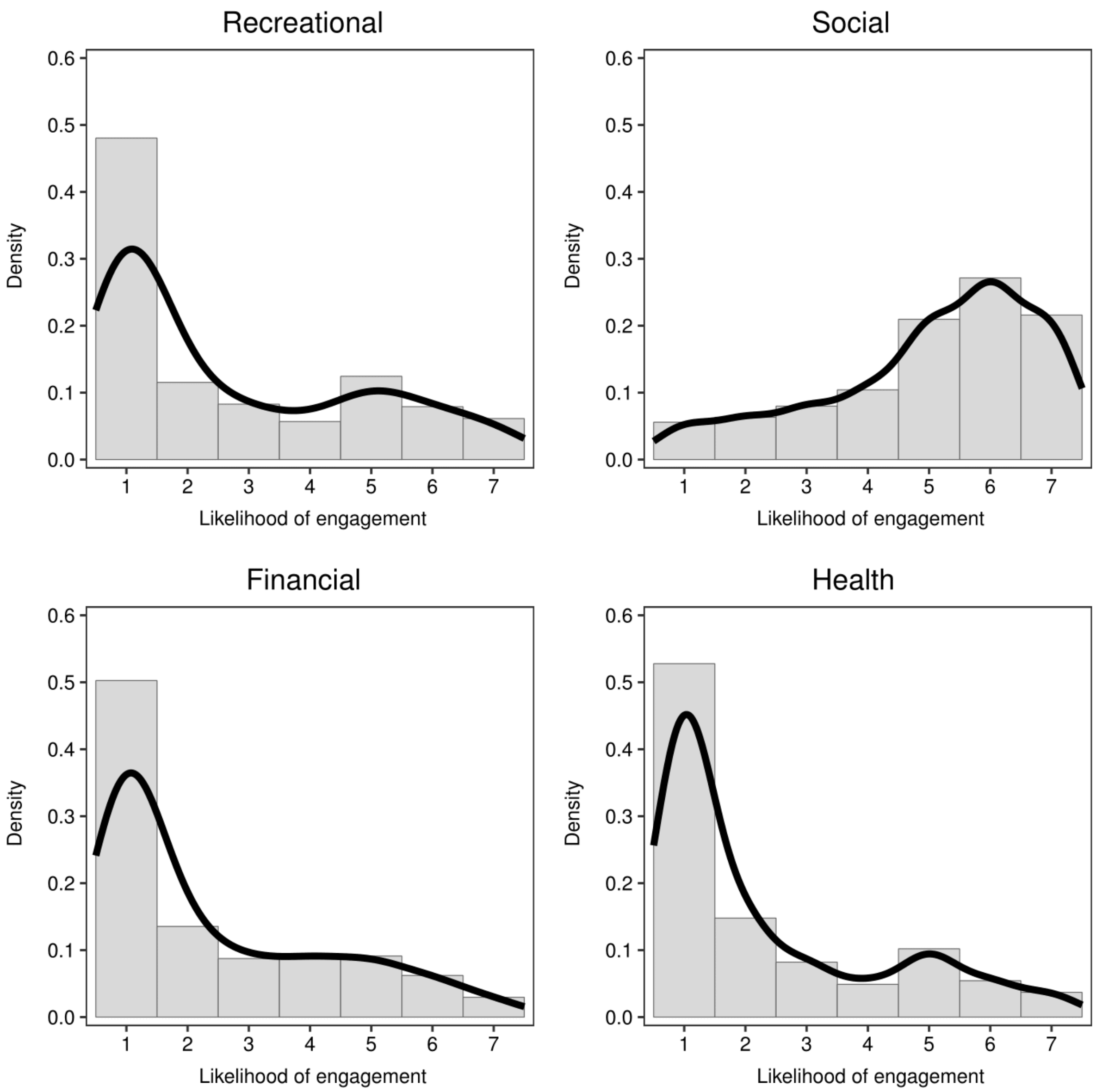

Figure 1. Ratings of 528 participants on the likelihood of engagement (i.e., risk-taking behavior) subscale of the DOSPERT in the recreational, social, financial, and health domains, reproduced from Rolison, Hanoch, Wood, and Pi-Ju (2014). The solid lines show the density (i.e., smoothed distribution) of participants' ratings across the 7-point Likert scale.

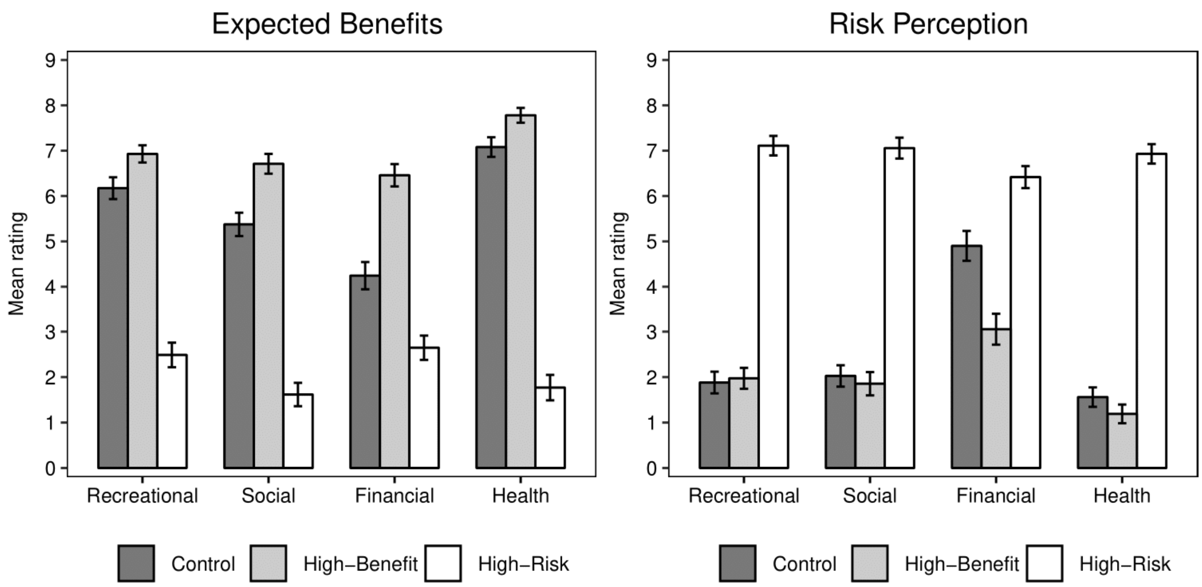

Figure 2. Study 1: Mean expected benefits and risk perceptions in each risk domain under control, high-benefit, and high-risk conditions.

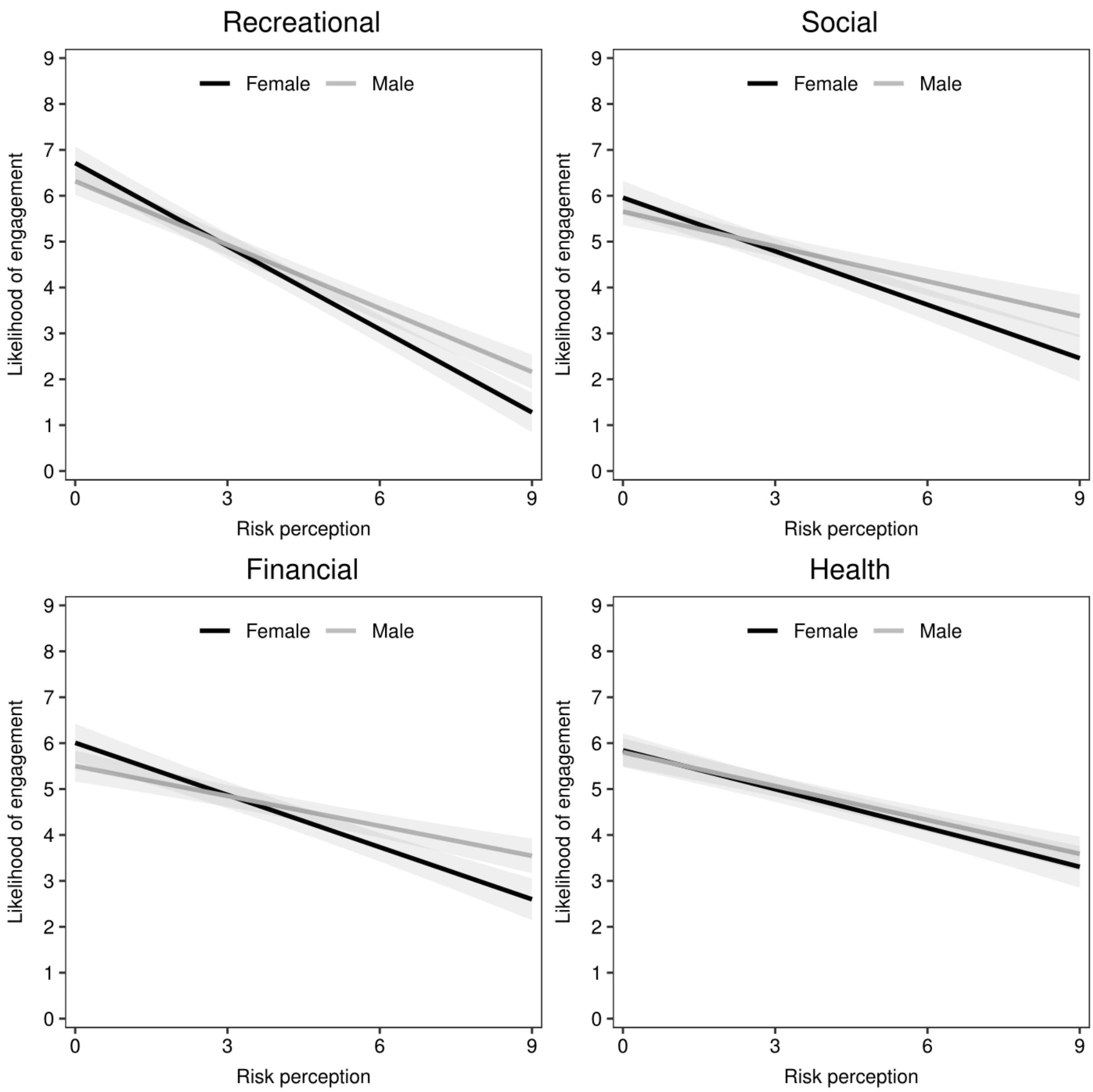

Figure 3. Study 2: Estimated likelihood of engagement for men and women in each risk domain, controlling for expected benefit ratings. The shaded areas indicate the 95% confidence intervals.

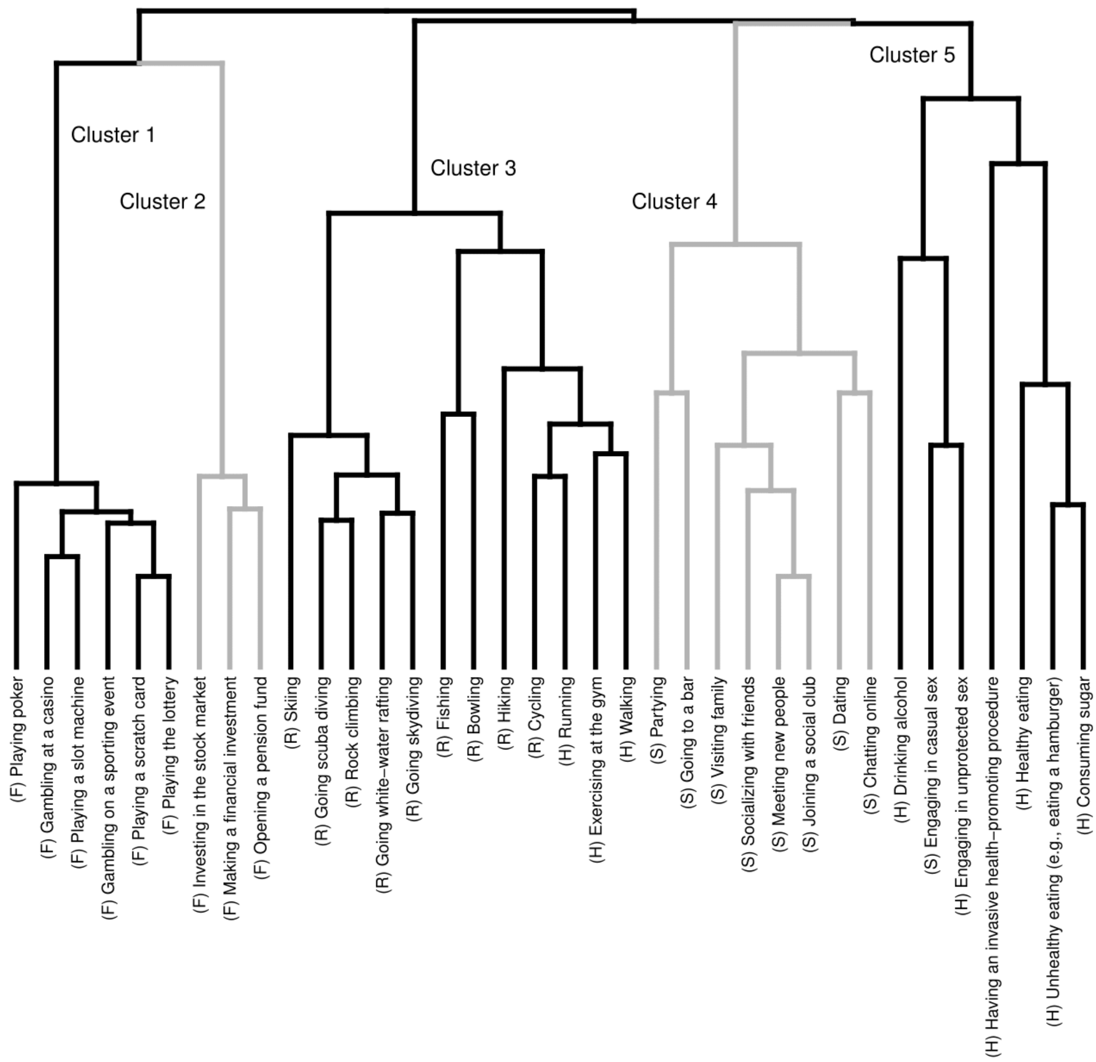

Figure 4. Study 3: Hierarchical cluster analysis on participants' judgements about the group of the scale items into domains.

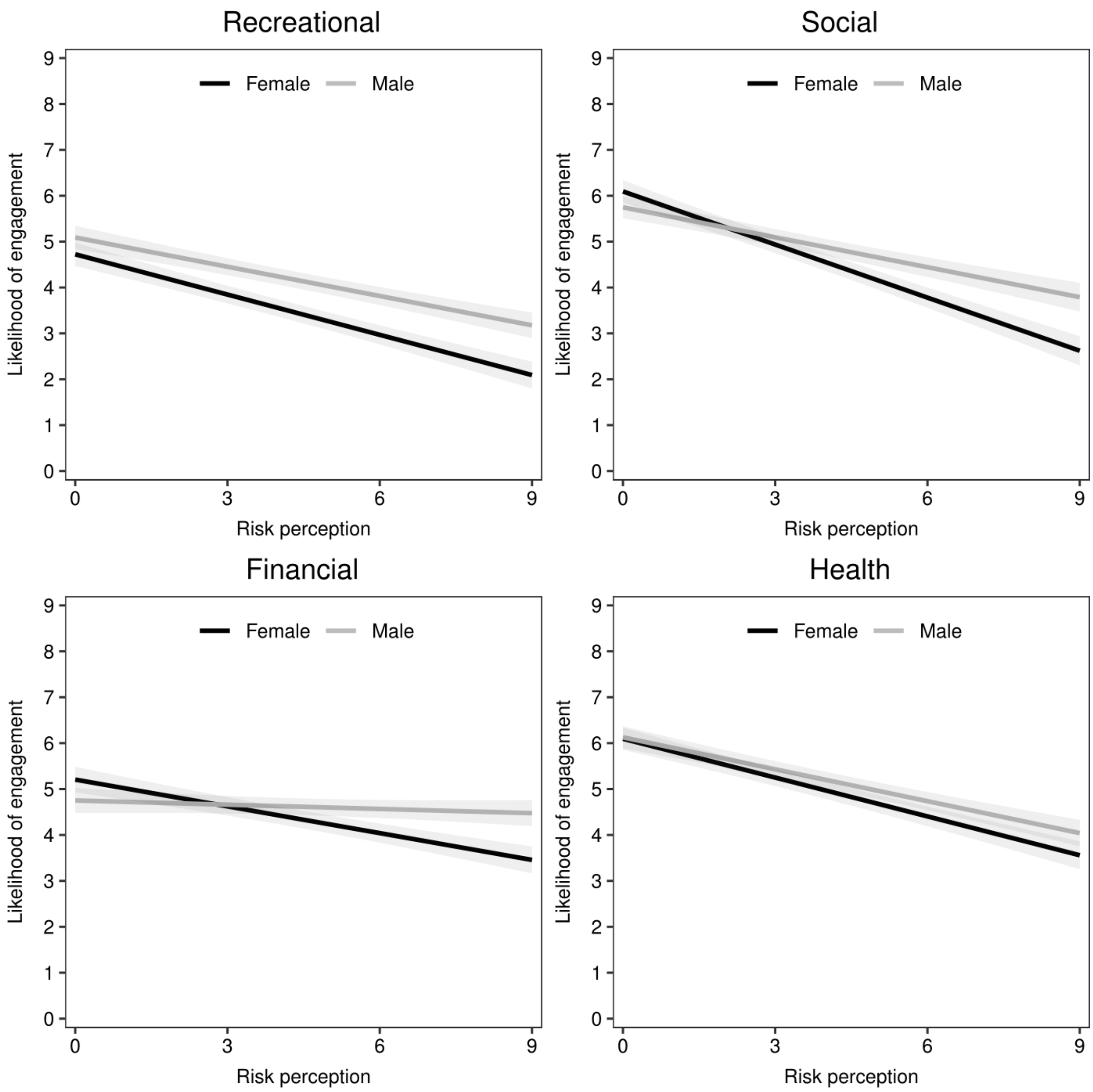

Figure 5. Study 3: Estimated likelihood of engagement for men and women in each risk domain, controlling for expected benefit ratings. The shaded areas indicate the 95% confidence intervals.

APPENDIX A

Table A1 provides the questionnaire items used in Study 2. The no. generated refers to the number of participants in the Study 1 sample who generated the item. The benefit and risk level refers to the level used to categorize items into low to high levels of expected benefits and risk perceptions. The modal response identifies the condition (control, high-benefit, and high-risk) generated the model number of responses for each item.

Table A1: Questionnaire items (Study 2)

Risk Domain	Questionnaire Item	No. generated	No. generated (men)	No. generated (women)	Benefit level	Risk level	Modal response (*n*)
Recreational	1. Going white-water rafting	1	1	0	High	High	High-risk (1)
	2. Hiking	18	7	11	High	Medium	High-benefit & Control (9)
	3. Reading	20	7	13	High	Low	Control (12)
	4. Going skydiving	11	5	6	Medium	High	High-risk (11)
	5. Skiing	9	3	6	Medium	Medium	Control (5)
	6. Watching television	10	6	4	Medium	Low	Control (7)
	7. Taking illicit drugs	13	5	8	Low	High	High-risk (12)
	8. Going scuba diving	2	1	1	Low	Medium	High-risk (2)
	9. Playing a video game	2	1	1	Low	Low	Control (2)
Social	1. Meeting new people	3	3	0	High	High	High-benefit (2)
	2. Socializing with friends	7	4	3	High	Medium	High-benefit (6)
	3. Visiting family	17	8	9	High	Low	High-benefit (10)
	4. Partying	8	5	3	Medium	High	High-benefit & High-risk (3)
	5. Dating	7	1	6	Medium	Medium	Control (4)
	6. Joining a social club	28	11	17	Medium	Low	High-benefit (15)
	7. Engaging in casual sex	12	1	11	Low	High	High-risk (12)
	8. Going to a bar	7	4	3	Low	Medium	High-risk (4)
	9. Chatting online	3	2	1	Low	Low	Control (2)

Financial	1. Investing in the stock market	14	8	6	High	High	High-benefit & Control (6)
	2. Making a financial investment	11	8	3	High	Medium	Control (6)
	3. Saving money	55	19	36	High	Low	High-benefit (37)
	4. Gambling at a casino	15	5	10	Medium	High	Control (10)
	5. Playing poker	16	8	8	Medium	Medium	High-risk & Control (6)
	6. Playing a scratch card	6	4	2	Medium	Low	Control (4)
	7. Playing a slot machine	16	7	9	Low	High	Control (12)
	8. Playing the lottery	10	3	7	Low	Medium	High-risk (6)
	9. Gambling on a sporting event	3	0	3	Low	Low	High-risk (3)
Health	1. Having an invasive health procedure (e.g., colonic irrigation)	1	1	0	High	High	High-risk (1)
	2. Exercising at the gym	19	7	11	High	Medium	Control (8)
	3. Healthy eating	52	12	40	High	Low	High-benefit (27)
	4. Engaging in unprotected sex	4	2	2	Medium	High	High-risk (3)
	5. Running	8	4	4	Medium	Medium	High-risk & Control (3)
	6. Walking	15	5	10	Medium	Low	Control (9)
	7. Smoking	33	12	21	Low	High	High-risk (32)
	8. Drinking alcohol	13	6	7	Low	Medium	High-risk (13)
	9. Unhealthy eating (e.g., eating a hamburger)	2	0	2	Low	Low	High-risk (2)

Table A2 provides the questionnaire items used in Study 3. The no. generated refers to the number of participants in the Study 1 sample who generated the item. The benefit and risk level refers to the level used to categorize items into low to high levels of expected benefits and risk perceptions.

Table A2: Questionnaire items (Study 3)

Risk Domain	Questionnaire Item	No. generated	No. generated (men)	No. generated (women)	Benefit level	Risk level
Recreational	1. Going white-water rafting	1	1	0	High	High
	2. Hiking	18	7	11	High	Medium
	3. Reading	20	7	13	High	Low
	4. Going skydiving	11	5	6	Medium	High
	5. Skiing	9	3	6	Medium	Medium
	6. Watching television	10	6	4	Medium	Low
	7. Taking illicit drugs	13	5	8	Low	High
	8. Going scuba diving	2	1	1	Low	Medium
	9. Playing a video game	2	1	1	Low	Low
	10. Cycling	9	2	7	High	Low
	11. Fishing	6	3	3	Medium	Low
	12. Rock climbing	6	1	5	Low	High
	13. Bowling	2	1	1	Low	Low
Social	1. Meeting new people	3	3	0	High	High
	2. Socializing with friends	7	4	3	High	Medium
	3. Visiting family	17	8	9	High	Low
	4. Partying	8	5	3	Medium	High
	5. Dating	7	1	6	Medium	Medium
	6. Joining a social club	28	11	17	Medium	Low
	7. Engaging in casual sex	12	1	11	Low	High
	8. Going to a bar	7	4	3	Low	Medium
	9. Chatting online	3	2	1	Low	Low
Financial	1. Investing in the stock market	14	8	6	High	High
	2. Making a financial investment	11	8	3	High	Medium
	3. Saving money	55	19	36	High	Low
	4. Gambling at a casino	15	5	10	Medium	High
	5. Playing poker	16	8	8	Medium	Medium
	6. Playing a scratch card	6	4	2	Medium	Low

	7. Playing a slot machine	16	7	9	Low	High
	8. Playing the lottery	10	3	7	Low	Medium
	9. Gambling on a sporting event	3	0	3	Low	Low
	10. Opening a pension fund	16	7	9	High	Low
Health	1. Having an invasive health procedure (e.g., colonic irrigation)	1	1	0	High	High
	2. Exercising at the gym	19	7	11	High	Medium
	3. Healthy eating	52	12	40	High	Low
	4. Engaging in unprotected sex	4	2	2	Medium	High
	5. Running	8	4	4	Medium	Medium
	6. Walking	15	5	10	Medium	Low
	7. Consuming sugar	5	1	4	Low	High
	8. Drinking alcohol	13	6	7	Low	Medium
	9. Unhealthy eating (e.g., eating a hamburger)	2	0	2	Low	Low

APPENDIX B

We conducted a repeated-measures analysis of variance (ANOVA) on participants' mean expected benefit and risk perception ratings across domains in Study 1. The categorized levels of expected benefit (low, medium, high) and risk perception (low, medium, high) were included as factors.

For expected benefit, the analysis yielded a significant effect of expected benefit category ($F(2,380) = 505.01$, $p < .001$, $\eta^2 = .73$), such that expected benefit ratings increased from low ($M = 3.05$) to medium ($M = 4.31$) to high ($M = 6.34$) expected benefit categories. There was also a significant effect of risk perception category ($F(2,380) = 244.26$, $p < .001$, $\eta^2 = .56$), whereby fewer benefits were expected as the risk perception category increased from low ($M = 5.14$) to medium ($M = 4.97$) to high ($M = 3.60$).

For risk perception, the analysis yielded a significant effect of risk perception category ($F(2,380) = 536.84$, $p < .001$, $\eta^2 = .74$), in which risk perception ratings increased from low ($M = 2.39$) to medium ($M = 3.83$) to high ($M = 5.60$) risk perception categories. There was also a significant effect of expected benefit category ($F(2,380) = 344.93$, $p < .001$, $\eta^2 = .65$), such that risks were perceived to be lower as the expected benefit category increased from low ($M = 4.85$) to medium ($M = 3.99$) to high ($M = 2.98$).

APPENDIX C

Figure 3C provides the sum of squared distances within clusters as a measure of the similarity of items within clusters (i.e., coherence within clusters) using the Elbow method, and the Silhouette coefficient as a measure of within-cluster consistency using the Silhouette method. Regarding the Elbow method, increasing the number of clusters from five (*SS* = 3.59) to six (*SS* = 3.00) clusters lead to relatively smaller reductions in the sum of squared within-cluster distances, compared to four clusters (*SS* = 4.86). Regarding the Silhouette method, within-cluster consistency was maximized by a five-cluster solution (coefficient = 0.50). Four-cluster (coefficient = 0.43) and six-cluster (coefficient = 0.49) solutions exhibited smaller Silhouette coefficients.

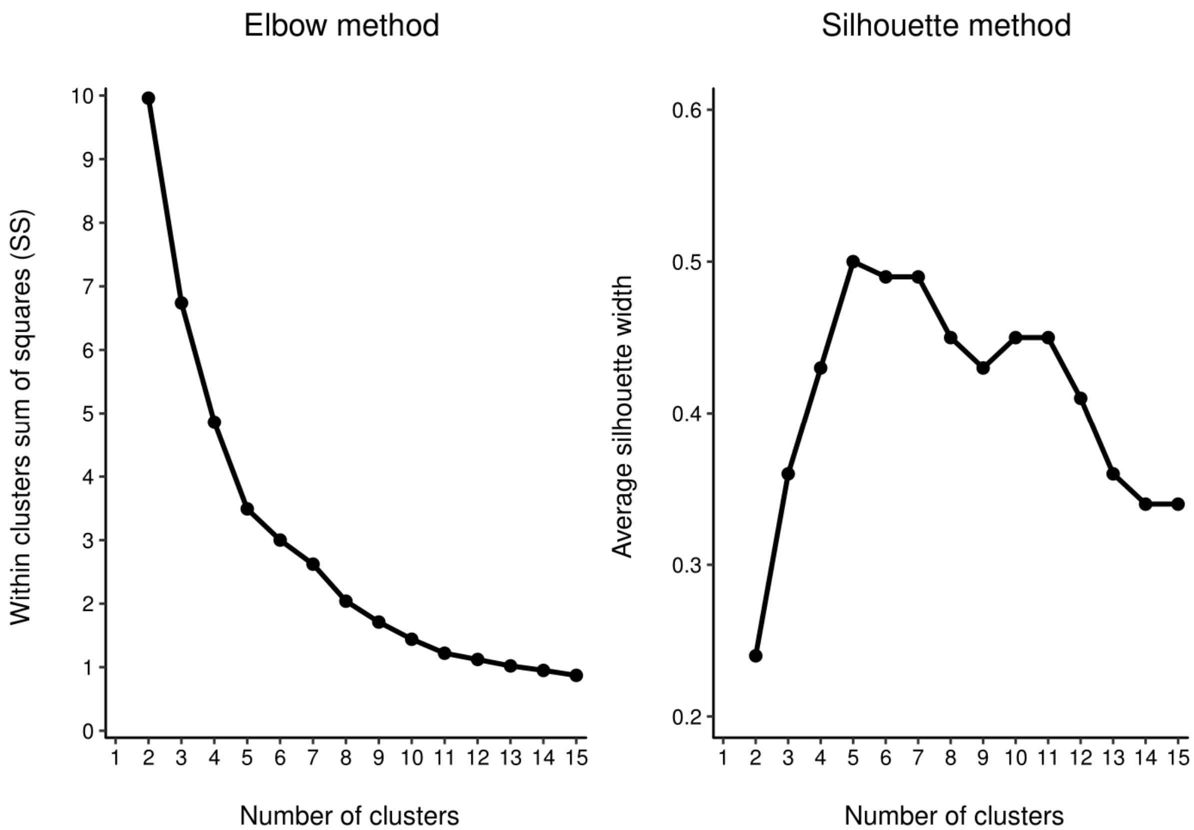

Figure C1. Study 3: Optimal number of clusters determined by the Elbow and Silhouette methods.