How Well Do We Know Our Inner Daredevil?

Probing the Relationship between Self-Report and Behavioral Measures of Risk Taking

Jonathan J. Rolison

Department of Psychology, University of Essex, UK

Thorsten Pachur

Max Planck Institute for Human Development, Berlin, Germany

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, signal detection theory, self-report, decision making, aging, ambiguity

Author note

Jonathan Rolison and Thorsten Pachur contributed equally to this work.

Abstract

To measure a person's risk-taking tendency, research has relied interchangeably on self-report scales (e.g., 'Indicate your likelihood of engaging in the risky behavior') and more direct measures, such as behavioral tasks (e.g., 'Do you accept or reject the risky option?'). It is currently unclear, however, how the two approaches map upon each other. We examined the relationship between self-report likelihood ratings for risky choice in a monetary gamble task and actual choice, and tested how the relationship is affected by task ambiguity (i.e., when part of the information about risks and benefits is missing) and age. Five hundred participants (aged 19-85 years) were presented with 27 gambles, either in an unambiguous or an ambiguous condition. In a likelihood rating task, participants rated for each gamble the likelihood that they would accept it. In a separate choice task, they were asked to either accept or reject each gamble. Analyses using a signal-detection approach showed that people's likelihood ratings discriminated between accept and reject cases in their choices rather well. However, task ambiguity weakened the association between likelihood ratings and choice. Further, older adults' likelihood ratings anticipated their choices more poorly than younger adults'. We discuss implications of these findings for existing approaches to the study of risk-taking propensity, which have often relied on self-reported risk tendency for ambiguous activities.

In the province of Quebec, some casino managers have made the remarkable step of allowing their clients to ban themselves from entering the establishment (Ladouceur, Jacques, Giroux, Ferland, & Leblond, 2002). Self-exclusion programs are intended to help gambling addicts avoid situations in which they believe they cannot resist temptation. Although many fail to comply with their agreement (Ladouceur et al., 2002; Ladouceur, Sylvain, & Gosselin, 2007), gamblers who commit to these programs do so because they anticipate that they will not be able to resist the lure of the casino. An ability to anticipate whether one will engage in a risky activity is crucial, as it empowers individuals, such as the self-excluding gamblers, to avoid situations in which their choices can have serious negative outcomes. Here, we ask how well people actually know the daredevil within them.

In psychology, researchers have employed various methodological approaches to assess individual differences in risk-taking tendency. One prominent approach has been to use self-report measures, where people are asked to indicate their likelihood to engage in a risky behavior (Blais & Weber, 2006; Rolison, Hanoch, Wood, & Pi-Ju, 2014; Weber, Blais, & Betz, 2002). For example, in the Domain Specific Risk Taking scale (DOSPERT; Weber et al., 2002) respondents are asked to evaluate their own likelihood of risk taking for various risky activities and behaviors (i.e., ‘Indicate your likelihood of engaging in…’) by rating themselves on a Likert scale (from 1 = ‘Not at all likely’ to 7 = ‘Extremely likely’). Individual differences in self-reported risk taking likelihood have been shown to be correlated with individual differences in real-world behaviors, such as the trading volume of financial investors (Markiewicz & Weber, 2013) and health behaviors, including smoking (Hanoch, Johnson, & Wilke, 2006).

However, self-report measures have potential shortcomings. For instance, individuals might lack insight into their own attitudes or behavioral tendencies and thus fail to accurately report on their likelihood of risk taking (Banaji, Hardin, & Rothman, 1993;

Greenwald & Banaji, 1995; Nisbett & Wilson, 1977). Individuals may also envision negative consequences of admitting to risky behaviors, motivating them to moderate their responses to comply with perceived social norms (Nederhof, 1985; Fisher, 1993). An alternative approach has been to measure behavior directly using decision making tasks (e.g., Bechara, Damasio, Tranel, & Damasio, 1997; Holt & Laury, 2002; Glöckner & Pachur, 2012; Figner, Mackinlay, Wilkening, & Weber, 2009; Wichary, Pachur, & Li, 2015). In these tasks, individuals decide on the basis of explicitly described or experienced outcomes and probabilities of the choice options. For example, respondents may be asked whether they accept a hypothetical gamble that offers a 25% chance to win $30 and a 75% chance to lose $10. Risk taking in such behavioral tasks has been shown to be associated with personality characteristics (Lauriola & Levin, 2001) and real world behaviors, such as smoking and drug use (Lejuez et al., 2002), and criminal offence (Pachur, Hanoch, & Gummerum, 2010; Rolison, Hanoch, & Gummerum, 2013).

An implicit assumption in this research is that self-reported likelihood of risk taking and actual choice behavior tap into the same underlying attitudes toward risk. In other words, if an individual takes few risks in their decision making, then they should also report a low likelihood of risk taking, indicating that they know their inner daredevil. On the other hand, studies on metacognition have revealed dissociations between self-judgments and behavior on a range of cognitive tasks (Koriat, 1997; Metcalfe, Schwartz, & Joaquim, 1993). For instance, people are often overconfident in the accuracy of their intuitive judgments and in their general knowledge (Griffin & Tversky, 1992; Koriat, Lichtenstein, & Fischhoff, 1980; but see Juslin, Winman, & Olsson, 2000). Further, people seem to have a limited ability to accurately predict the impact of outcome magnitudes and probabilities of options on their actual choice (e.g., Morewedge, Gilbert, Keysar, Berkovits, & Wilson, 2007; Gilbert, Morewedge, Risen, & Wilson, 2004). In studies of memory, subjective confidence and actual recall accuracy are often poorly correlated (e.g., Bothwell, Deffenbacher, & Brigham, 1987). One reason is that when asked to rate how confident they are in memory recall, people tend to consider in their ratings also factors that they believe do but in fact do not improve memory (e.g., luminance; Busey, Tunnicliff, Loftus, & Loftus, 2000; Rhodes & Castel, 2008). In Rhodes and Castel (2008), participants predicted that they would better recall words presented in a larger font size, despite font size having little actual effect on recall. People have also been shown to express different preferences among options depending on whether the preference is elicited through a behavioral choice or a rating task (Goldstein & Einhorn, 1987; Lichtenstein & Slovic, 2006). Despite these reasons for possible discrepancies between self-ratings of risk taking likelihood and actual choice behavior, to our knowledge no previous study has explored how the two measures of risk propensity map upon each other.

Our goal in this article is to fill this gap. To characterize the relationship between self-report ratings of the likelihood of taking a risk and actual risky choice, we use a signal detection theory (SDT) approach (Green & Swets, 1966; Macmillan & Creelman, 2005). From this perspective, a likelihood rating is seen as an attempt to discriminate between cases where a gamble is accepted (signal trial) and cases where a gamble is rejected (noise trial). Because gambles vary in their attractiveness and because choice behavior is stochastic (e.g., Mosteller & Nogee, 1951), acceptance and rejection cases are represented as two probability distributions. One end of the continuum represents a low attractiveness of a gamble, whereas the other end represents a high attractiveness. To the extent that an individual's likelihood ratings accurately discriminate between acceptance and rejection cases, the overlap between the two distributions is larger or smaller. For example, if the two distributions do not overlap at all, then high likelihood ratings are given only to those cases where a gamble is accepted. On the other hand, if the two distributions overlap entirely, then the likelihood ratings are entirely dissociated from the person's actual choice behavior.

The SDT framework is useful because it allows us to disentangle discriminability (or sensitivity) and response criterion in the likelihood ratings. Discriminability represents the accuracy with which acceptance and rejection cases can be told apart; the response criterion, in contrast, represents the threshold on the strength of attractiveness continuum beyond which gambles receive a high likelihood of being chosen (i.e., higher than the midpoint of the scale, representing neither likely nor unlikely). For example, with a high (i.e., conservative) response criterion only few cases receive a high likelihood rating. Conversely, with a low (i.e., liberal) response criterion many cases receive a high likelihood rating. The SDT framework thus enables us to independently assess how sensitively likelihood ratings reflect actual choice behavior as well as identify response tendencies in the likelihood ratings. For instance, it could be that in the likelihood ratings respondents have a bias to downplay their risk-taking tendency, which would be indicated by a conservative threshold. As we describe in the next two sections, with the SDT measures of discriminability and response criterion we can also test how the mapping between likelihood ratings and actual choice is affected by the ambiguity of the options in the task and age, and to what extent people's response tendencies in the likelihood rating task are adaptive—in the sense that they respond to differences in the frequency of risk-seeking behavior.

Does Ambiguity Affect the Correspondence between Self-Report and Choice?

Laboratory tasks that measure risk taking typically provide complete information about all the outcomes (e.g., win $30 or lose $10) and probabilities (e.g., 25% chance to win and 75% chance to lose) of the choice options. In many real world situations, however, decisions must be made without the luxury of knowing all the possible outcomes and probabilities. For example, the chance of winning a national lottery jackpot rise and fall according to weakly ticket sales; and a homebuyer cannot know how financial markets will influence their future mortgage repayments. Many of the activities used in self-report risk taking scales represent such ambiguous options. For instance, one may not know all the possible consequences associated with ‘Going white-water rafting at high water in the spring’ or ‘Betting a day’s income at a high-stake poker game’ (DOSPERT; Blais & Weber, 2006), let alone their precise probabilities. Ambiguity likely imposes additional demands on people’s ability to self-reflect on the likelihood of their risk behavior. When activities are vague about their possible outcomes (e.g., ‘Going white-water rafting at high water in the spring’; DOSPERT; Blais & Weber, 2006), or the possible outcomes and probabilities are unknown and need to be inferred or estimated, people must engage greater cognitive effort to assess their own likelihood of risk taking. If ambiguity weakens the degree to which self-reported likelihood ratings and choice behavior are associated, then this would have implications for the reliability of self-report scales.

Further, people are less likely to choose an option when some of its characteristics are ambiguous (i.e., one or more of the outcomes or probabilities is unknown; Ellsberg, 1961; see also Camerer & Weber, 1992; Hsu et al., 2005) than when all characteristics are known. Does people’s criterion setting in the likelihood ratings reflect this difference in choice? Analyses of criterion setting in discrimination tasks have shown that people adaptively adjust their response criterion according to the base rate of signal events (Estes & Maddox, 1995; Rhodes & Jacoby, 2007). For example, in a memory study, Estes and Maddox (1995) found that participants shifted to a more liberal criterion when memory test sets contained a majority of previously studied (i.e., old) items compared to when the proportion of old and new cases was balanced. Are a person’s likelihood ratings of risky choices similarly sensitive to the reduced tendency to choose a risky option under ambiguity? If so, people’s response criterion should be more conservative than when an option’s outcomes and probability are fully provided.

Reduced Correspondence Between Self-Report and Choice in Older Adults?

To the extent that, as described above, accurate likelihood ratings require greater reflective effort than choices, discriminability may be reduced in older than in younger adults. Controlled cognitive processes (e.g., explicit memory) that are linked to metacognitive abilities necessary for self-reflection show age-related decline (Hartshorne & Germine, 2015; Salthouse, 2006). Moreover, relative to younger adults older adults seem to be constrained in drawing samples from memory (Hansson, Rönnlund, Juslin, & Nilsson, 2008)—which might be necessary to accurately assess the likelihood of one's behavior. Older adults also show greater decrements in decision quality when choosing between multiple options than when choosing between only two (Frey, Mata, & Hertwig, 2015). Hence, older adults may be poorer than younger adults at discriminating risky and safe choices on the basis of their likelihood ratings. If so, age-related differences on self-report measures of risk taking may be biased by age differences in people's ability to self-reflect on their choice behavior.

Further, a wealth of research exploring individual differences in risk taking has shown that older adults are typically less willing to take risks than younger adults (Denburg, Tranel, & Bechara, 2005; Henninger, Madden, & Huettel, 2010; Rolison, Hanoch, & Wood, 2012; Zamarian et al., 2008). If response criterion in likelihood ratings is adaptive, older adults should show a more conservative criterion.

Aims of the Current Study

To examine the relationship between self-reported likelihood of choosing a risky option and actual choice behavior, participants were shown the same set of gambles in two types of tasks. In one of the tasks, they were asked to report their likelihood of risk taking ("Indicate the likelihood that you would accept this gamble"), and in the other task, to make choices ("Do you accept or reject this gamble?"). On the basis that self-report measures typically study ambiguous real world activities whereas behavioral tasks usually make information about all possible outcomes available, we examined whether task ambiguity affects the relationship between likelihood ratings and choice. On the basis that likelihood ratings might require greater reflective effort than choices, we further examined whether individual differences in decision making, and in particular age differences, affect the mapping between likelihood ratings and choice behavior. In addition, we tested whether reductions in the willingness to choose a risky option under ambiguity and in older adults would be accompanied by a corresponding shift in response criterion in the likelihood ratings; and whether individual differences in risky choice in general are accompanied by differences in response criterion.

How Well Do People Know Their Inner Daredevil?

Methods

Participants. We recruited N = 500 respondent (245 females) via Mechanical Turk on Amazon. Data reliability of the Amazon Mechanical Turk participant pool has been validated elsewhere by comparison with other recruitment methods (Berinsky, Huber, & Lenz, 2012; Paolacci, Chandler, & Ipeirotis, 2010). Participants were awarded $1.00 on completion of the unambiguous task and $1.50 on completion of the ambiguous task owing to the extended length of the task (see Materials and Procedure). Fifteen participants in the unambiguous condition and 41 participants in the ambiguous condition failed to complete the study and were thus removed from all our analyses to follow. All were United States (US) residents. Participants' internet protocol (IP) address was used to confirm their geolocation in the US. Participants took on average 13 minutes and 40 seconds (SD = 7 minutes: 21 seconds) to complete the study. Participants ranged from 19 to 85 years of age (M = 44.86; SD = 15.71). One hundred twenty five were aged 19-30 years, 102 were aged 31-40 years, 71 were aged 41-50 years, 77 were aged 51-60 years, 119 were aged 61-69 years, and six were aged 70-85 years. Almost all participants (98%) had completed lower secondary or vocational education and more than half (62%) had completed higher vocational or university education. A minority (7%) had an annual household income below $10,000. For most, their household income ranged $10,000 and $50,000 (54%) or $50,000 and $60,000 (30%). Few (9%) had a household income above $100,000.

Materials. We constructed 27 two-outcome gambles using a factorial design (see Appendix A), each consisting of a gain amount ($10, $20, $30), a loss amount ($10, $20, $30), and chances to win and lose, respectively (25%, 50%, 75%). In the unambiguous condition, complete information about the gain and loss amounts and the chances to win and lose of each gamble was provided. In the ambiguous condition, the gain amount, loss amount, or the chances to win and lose was not provided (as indicated by a "?"; Appendix A).

Design and procedure. Participants were randomly assigned to receive either unambiguous (N = 249) or ambiguous (N = 251) gambles (see Appendix B for instructions). In a *likelihood rating task*, participants viewed the same 27 gambles and were asked "Please indicate the likelihood that you would accept this gamble" on a 7-point scale (1 = "extremely unlikely", 2 = "moderately unlikely", 3 = "somewhat unlikely", 4 = "not sure", 5 = "somewhat likely", 6 = "moderately likely, 7 = "extremely likely"). The likelihood rating scale was modelled after rating scales used in the literature to measure risk-taking propensity (Blais & Weber, 2006; Weber et al., 2002). In a *choice task*, participants were asked for each of the 27 gambles "Do you accept or reject this gamble?". They indicated choice by selecting an "accept" or "reject" option. The order of the two tasks was counterbalanced. Within each task, participants were presented each gamble one at a time in random order. A blank screen followed each response before presentation of the next gamble in the set.

In the ambiguous condition, participants were additionally presented with a third task that followed the choice task and likelihood rating task and were asked to indicate for each of the 27 gambles what they believed to be the unknown gamble amounts and chances. We recorded participants' responses about the missing values on the basis that their judgments might provide an indication of their perceptions of the expected value of ambiguous gambles (see Appendix C for full description). The gambling problems were presented in the same format as in the choice and likelihood rating tasks. Participants were asked “What do you think is the most likely amount that can be [won, lost]” for ambiguous gambles in which the gain or loss amount was unknown, and were asked “What do you think are the most likely chances of winning and losing” when the outcome chances were unknown. Participants chose among the candidate amounts and chances. Finally, all participants then provided their demographic information.

Results

We first briefly summarise analyses of participants’ choices and likelihood ratings (see Appendix C for full description), showing that they exhibit several established regularities. Specifically, participants accepted gambles with a higher expected value more often than ones with a lower expected value and they also provided higher likelihood ratings for the former. Further, gambles were less often accepted in the ambiguous than in the unambiguous condition and there was a trend toward lower likelihood ratings for ambiguous gambles; replicating previous findings, participants were thus ambiguity averse. We also found that compared to younger participants, older participants accepted the gambles less frequently and also provided lower likelihood ratings. Two-way interactions revealed that participants were less responsive to differences in the expected value of ambiguous gambles than they were for unambiguous gambles in both their decisions and likelihood ratings. Age interacted with the expected value of gambles, such that older age was associated with reduced sensitivity to differences in the gambles’ expected values.

Next, we examined the relationship between participants’ likelihood ratings and choices by conducting an SDT analysis. Hit and false alarm rates were calculated individually for each participant. Hit rates equalled the total number of accepted (in the choice task) gambles where the participant indicated (in the likelihood rating task) a high likelihood rating (i.e., > 4) and half of the cases with a neutral rating (i.e., = 4) divided by the total number of accepted gambles. False alarm rates equalled the total number of rejected (in the choice task) gambles where the participant indicated (in the likelihood rating task) a high likelihood rating (i.e., > 4) and half of the cases with a neutral rating (i.e., = 4) divided by the total number or rejected gambles. To ensure robust hit and false alarm rates also when there are only few signal and noise trials, in the calculation of the hit and false alarm rates 0.5 was added to the numerator and 1 to the denominator (Snodgrass & Corwin, 1988). Discriminability scores, *d'*, were calculated as the standardized difference between the hit and false alarm rates ($d' = z[\text{hit}] - z[\text{false alarm}]$) and provide a measure of how well participants' likelihood ratings discriminated between choices to accept and reject a gamble. A score of 0 indicates that a participant's likelihood ratings do not discriminate between their accepted and rejected gambles and scores > 0 indicate better discriminability. Response criterion scores, *C*, were calculated as the mean of the standardized hit and false alarm rates ($C = -0.5 \times [z\{\text{hit}\} + z\{\text{false alarm}\}]$) and provide a gauge to the threshold on the attractiveness dimension past which gambles are given a high likelihood rating. Positive scores represent a conservative criterion; negative scores represent a liberal criterion.

Our SDT analysis showed that whether or not the choice task was completed before or after the likelihood rating task had no significant influence on discriminability ($M_{\text{choices first}} = 1.90$, $M_{likelihood\ ratings\ first} = 1.78$; $t(498) = 1.55$, $p = .123$) or the response criterion ($M_{\text{choices first}} = 0.23$, $M_{likelihood\ ratings\ first} = 0.15$; $t(498) = 1.72$, $p = .086$), indicating that participants' likelihood ratings did not simply accord better with their choices when they had already completed the choice task. Figure 1A shows the average likelihood ratings as a function of the percentage of accepted gambles split at low and high levels of discriminability. As can be seen, for participants with lower discriminability the average likelihood ratings were slightly more regressive, and therefore less indicative of the proportion of accepted gambles, than for participants with higher discriminability. Figure 1B shows the likelihood ratings as a function of the percentage of accepted gambles separately for participants who were liberal (i.e., response criterion < 0) and conservative (i.e., response criterion > 0) in their ratings on the likelihood scale. Liberal participants awarded higher likelihood ratings to gambles than did conservative participants (Figure 1B).

Did ambiguity affect the correspondence between likelihood ratings and choice, and if so, how? As shown in Figure 2A, participants exhibited lower discriminability in the ambiguous (M = 1.60, SD = 0.82) than in the unambiguous condition (M = 2.09, SD = 0.89). This difference was confirmed by an independent-samples t-test ($t(498) = -6.33$, $p < .001$) and implies that participants' likelihood ratings discriminated between acceptance and rejection cases less accurately in the ambiguous than in the unambiguous condition. The poorer discriminability for ambiguous gambles may have resulted simply from greater inconsistency in participants' ambiguous gamble choices. To assess the role of choice consistency on discriminability, we took advantage of the nine 3-item sets of gambles in the stimulus set for which two of the attributes were identical and the third varied (see Appendix A). For example, for one set of three gambles, the gain amount was equal to $10, the chances to win and lose were equal to 25% and 75%, respectively, and the loss amount increased from $10, $20, and $30, respectively. As a measure of consistency, we determined whether participants showed a monotonic choice pattern across the items as the loss amount increased from $10 to $30. For instance, accepting the $10 and the $30 losses, but rejecting the $20 loss, or rejecting the $10 and $30 losses, but accepting the $20 loss would indicate inconsistent choice behavior. For each participant, we counted for how many of the nine sets of gambles they showed a consistent choice pattern. Not surprisingly, participants were less consistent for ambiguous gambles (M = 95%, SD = 0.11) than for unambiguous gambles (M

= 97%, SD = 0.07; group difference, $t(498) = 2.74$, $p = .006$). However, when controlling for choice consistency in an analysis of covariance (ANCOVA), participants still exhibited poorer discriminability in the ambiguous ($M_{marginal} = 1.65$) than in the unambiguous condition ($M_{marginal} = 2.04$; $F(1,497) = 32.14$, $p < .001$).

Also response criterion differed between the ambiguous and unambiguous conditions, with the criterion being more conservative in the former than in the latter (Figure 2B; $M_{ambiguous} = 0.24$, $SD = 0.60$; $M_{unambiguous} = 0.14$, $SD = 0.49$; $t(498) = 2.14$, $p = .033$; Panel C in Figure A1). However, as reported earlier, participants were also less likely to accept the gamble in the ambiguous than in the unambiguous condition. Controlling risk taking—measured as the percentage of accepted gambles—in an ANCOVA, differences in response criterion between the ambiguous ($M_{marginal} = 0.19$) and unambiguous conditions ($M_{marginal}$ = 0.18) disappeared ($F(1,497) = 0.07$, $p = .786$). This indicates that the differences in response criterion between the ambiguous and unambiguous condition largely reflected an adaptive response to the differences in the frequency of acceptance cases between the conditions.

How does age influence the correspondence between likelihood ratings and choice? There was a quadratic age trend in discriminability in the ambiguous condition ($\beta_{linear} = 1.82$, $t = 3.70$, $p < .001$; $\beta_{quadratic} = -1.77$, $t = 3.62$, $p < .001$), but no age effect in the unambiguous condition ($\beta_{linear} = .78$, $t = 1.65$, $p = .100$; $\beta_{quadratic} = -.73$, $t = 1.55$, $p = .123$). However, when controlling for choice consistency, there was a significant quadratic age trend in discriminability in both the unambiguous ($\beta_{linear} = 1.11$, $t = 2.64$, $p = .009$; $\beta_{quadratic} = -1.03$, t = 2.45, $p = .015$) and ambiguous condition ($\beta_{linear} = 1.05$, $t = 2.48$, $p = .014$; $\beta_{quadratic} = -0.99$, t = 2.32, $p = .021$). Probing the estimated slopes, discriminability changed little from age 19 years ($d'_{unambiguous} = 1.56$; $d'_{ambiguous} = 1.09$) to age 40 years ($d'_{unambiguous} = 1.75$; $d'_{ambiguous}$ = 0.98), whereupon it reduced sharply with age by age 60 years ($d'_{unambiguous} = 1.11$; $d'_{ambiguous}$ = 0.06) to 70 years ($d'_{unambiguous} = 0.49$; $d'_{ambiguous} = -0.70$) and into older age (80 years,

$d'_{unambiguous}$ = -0.33; $d'_{ambiguous}$ = -1.66). Similarly, a linear regression analysis revealed a quadratic age trend in response criterion in the ambiguous condition (β_{linear} = 1.22, t = 2.46, p = .015; $\beta_{quadratic}$ = -1.21, t = 2.43, p = .016), but no age effect in the unambiguous condition (β_{linear} = -.24, t = 0.50, p = .617; $\beta_{quadratic}$ = .30, t = 0.63, p = .530).

Finally, we tested for a general association between individual differences in risk taking—measured as the percentage of accepted gambles—and the response criterion. A linear regression revealed that higher risk taking was strongly associated with a lower response criterion (β = -.62, t = 17.48, p < .001; Figure 2D). This strong association remained after controlling for individual differences in age (β = -.62, t = 17.44, p < .001), which we found previously were correlated with the response criterion. Inspecting Figure 2D, participants who accepted fewer than half of the gambles (i.e., were risk averse) had a conservative response criterion, which means that they falsely identified few instances in which they rejected a gamble (low false alarm rate), but also missed many instances in which they accepted a gamble (low hit rate rate). Conversely, participants who accepted more than half of the gambles (i.e., were risk seeking) had a liberal response criterion (Figure 2D), meaning that in their likelihood ratings they falsely identified many instances in which they rejected a gamble (high false alarm rate), but also identified many instances in which gambles were accepted (high hit rate). This again is indicative of adaptive response criterion setting. Additionally, and surprisingly, risk taking was also associated with discriminability (β = -.12, t = 2.70, p = .007): participants who accepted a higher number of gambles tended to show lower discriminability (Figure 2C).

Discussion

Research on individual differences in risk taking has implicitly assumed that people have a very good sense of their inner daredevil and have thus used direct, behavioral tasks and self-report measures of risk-taking propensity more or less interchangeably. Here, we investigated how closely self-reported likelihood of risk taking agrees with actual choice behavior.

Items in self-report risk taking scales typically refer to rather ambiguous activities (e.g., ‘Betting a day’s income at a high-stake poker game’: DOSPERT; Blais & Weber, 2006), in which one or more of the outcomes or their probabilities is unknown. In behavioral tasks, on the other hand, complete information about all possible choice options is usually either provided (Holt & Laury, 2002; Figner et al., 2009) or this information can be learned over the course of the experimental session (Bechara et al., 1997; Lejuez et al., 2002). We therefore tested whether the mapping of likelihood ratings onto choices is affected by whether the options are ambiguous or not and found that participants were less able to discriminate between accept and reject decisions for ambiguous problems than for unambiguous problems (Figure 2). The poorer discriminability also held after controlling for greater intra-individual inconsistency in participants’ choices for ambiguous gambles. It thus appears that when choice options are ambiguous, likelihood ratings do not reflect the risk taking tendencies that determine choice behavior as closely as when the options are unambiguous. One possible reason could be that ambiguity imposes additional cognitive demands. When the possible consequences of a risky activity are vague (e.g., ‘Going white-water rafting at high water in the spring’; DOSPERT; Blais & Weber, 2006) or when one or more of the possible outcomes or probabilities is unknown (e.g., ‘Betting a day’s income at a high-stake poker game’) the task requires mental simulation of the unknown possible outcomes and probabilities. In situations of ambiguity, people might, for instance, infer lower chances from larger payoffs (Pleskac & Hertwig, 2014). When reflecting on the likelihood of their choice behavior, the mental simulation required in situations of ambiguity presents an additional challenge for anticipating one’s choices.

The finding that ambiguity weakens the association between likelihood ratings and choice has implications for the construction of self-report scales. Researchers need to be cognizant that items that are highly ambiguous (e.g., ‘piloting a small plane’; DOSPERT; Blais & Weber, 2006) may not very closely relate to actual risk taking and distort findings based on self-report scales. From this perspective, it could be interesting to consider to what extent observed domain differences in self-reported risk taking may be due, in part, to differences in the degree of ambiguity in some of the items (e.g., Health; ‘Drinking heavily at a social function; DOSPERT; Blais & Weber, 2006). To examine this, one could attempt to avoid overly ambiguous items and strive to hold the level ambiguity constant across the sets of items. Some items could be made less ambiguous by including more explicit information about the possible outcomes and probabilities. For example, items such as ‘Betting a day’s income on the outcome of a sporting event’ could be modified to reduce ambiguity and read: ‘Betting a day’s income on the outcome of a sporting event when the chances of winning are advertised to be 1 in 5’. Further, ‘Going down a ski run that is beyond your ability’ could be modified to ‘Going down a ski run that is beyond your ability and break your leg with a chance of 10%’. To foster comprehension of the added risk information (and to avoid that responses are driven more by the person’s numeracy than their risk propensity), the information could be presented in a graphical format, such as icon arrays (Rolison, Morsanyi, O’Connor, 2015). Nevertheless, note that in several real-world domains risk information is naturally present in numerical format (e.g., the betting odds at a sporting event).

Another key finding was that discriminability was negatively affected by age, and this also held when controlling for age differences in choice consistency. It thus seems that older adults are less able to anticipate their own risky choices when asked to do so on a likelihood rating scale. Our finding appears to be at odds with research that has shown a stronger association between intention and behavior with advancing age (Downs &

Hausenblas, 2005; Hagger, Chatzisarantis, & Biddle, 2002). However, behaviors used to study the intention-behavior gap—such as physical exercise, quitting smoking, eating behavior, and alcohol use—tap into goal setting and implementation, to which older adults may have more experience than younger adults (e.g., Hagger et al., 2002). One possible explanation for our finding is age-related decline in controlled cognitive processes (e.g., Hartshorne & Germine, 2015). Self-report, multi-option likelihood scales might require more cognitive effort than do simple, binary choice tasks (cf. Frey et al., 2015). Further, the age differences in discriminability might be due to differences in the ability to mentally simulate, based on episodic samples drawn from memory, scenarios that are necessary to accurately gauge the likelihood of one's own future behavior. Hansson, Rönnlund, Juslin, and Nilsson (2008), for instance, concluded that older adults' ability to draw samples from memory is reduced, hampering their accuracy in metacognitive confidence judgments (see also Hansson, Juslin, & Winman, 2008). As we did not assess participants' cognitive functioning, however, we can only speculate about the possible reasons for age-related reduction in discriminability in risky choice behavior. Further research could explore to what extent it is indeed lower cognitive ability that weakens the mapping between likelihood ratings and choice behavior. This avenue of research could reveal new insight into the degree to which some risk taking measures are more demanding than others and whether a minimum level of cognitive ability may be necessary for reliable responding. Nevertheless, our findings imply that researchers need to be careful when drawing inferences about age differences in risk taking irrespective of the type of measure used.

Unexpectedly, we found that discriminability was also affected by risk taking tendency, such that likelihood ratings were less discriminative of risky choices among participants who accepted many gambles than among those who accepted only few gambles. This finding could imply that groups of individuals who more often engage in risky activities (e.g., offenders; Pachur et al., 2010; Rolison et al., 2013) may be less able to report reliably on their likelihood of risk taking. Hence, group differences in apparent risk taking might depend on whether risk taking propensities are assessed using self-report or behavioral tasks. Nevertheless, further research is required to establish whether this finding can be replicated.

Likelihood rating scales and behavioral tasks differ both in their reliance on self-report and in their response format. Self-report tasks typically use Likert scales, whereas binary choice options are normally used to elicit preferences in behavioral tasks. A wealth of research has shown that expressed preference can differ as a function of how it is elicited (Goldstein & Einhorn, 1987; Lichtenstein & Slovic, 2006). For instance, while people may choose a small reward that is likely over a larger reward that is less likely, they will often assign the latter a higher numerical value (Lichtenstein & Slovic, 2006). Our comparison of self-reported likelihood of risk taking and choice behavior confounds effects self-report with the influence of response format. Crucially, however, our current motivation was to assess the degree to which likelihood ratings used in self-report tasks map upon actual choice behavior. Prominent risk taking questionnaires (e.g., DOSPERT, Weber et al., 2002; gambles, Holt & Laury, 2002) regularly used to elicit risk preferences equally confound self-report and response format. Future research that seeks to disentangle these two features of risk taking measures could promote the development of risk taking scales that afford a better mapping across measures.

Despite the above constraints on the correspondence between likelihood ratings and choice, overall the former seems to be a good proxy for the latter. Moreover, our data show in several ways that in their likelihood rating participants set their response criterion—the threshold on the strength of attractiveness continuum beyond which gambles receive a high likelihood of being accepted—adaptively to the base rate of their risky choices. Specifically differences in acceptance of the gamble between ambiguous and unambiguous conditions, due to age, and individual differences in general were accompanied by parallel shifts in response criterion: when people made more risky choices, they also tended to set a more liberal criterion in their likelihood ratings. This finding dovetails with results from studies of recognition memory, showing that people adaptively adjust their response criterion according to the base rate of signal events (Estes & Maddox, 1995; Rhodes & Jacoby, 2007).

Our study has a number of possible limitations. First, we studied people's choice behavior for gambles with hypothetical outcomes, rather than ones that had real financial consequences. However, our current goal was to probe the relationship between self-report and choice for tasks that had similar potential outcomes. Had we incentivised responses in the choice task, the consequences of participants' choices could have distorted any natural relationship between likelihood ratings and choice. Further research may seek to explore whether this relationship is strengthened or weakened when both the likelihood ratings and choices are incentivised by real financial outcomes. Second, we asked participants to report on their likelihood of accepting monetary gambles. Self-report scales are designed to capture broad behavioral tendencies that are stable across occasions and situations (e.g., 'Driving a car without wearing a seatbelt', DOSPERT; Blais & Weber, 2006). The task we presented to participants, especially in the context of unambiguous gambles, contained highly specific one-shot instances for which self-reported risk taking may be more variable across occasions and situations. Nonetheless, choices between specific monetary gambles are commonly used to estimate individuals' underlying risk attitude (Becker, Deckers, Dohmen, Falk, & Kosse, 2012; Dohmen, Falk, Huffman, & Sunde, 2010; Glöckner & Pachur, 2012; MacCrimmon & Wehrung, 1990). Third, we studied the relationship between self-report and choice only in the financial domain. While it was important to ensure for our present purposes that participants' likelihood ratings and their choices were both based on the same gambling problems, the relationship between self-report and choice might depend on the risk domain. Potentially, some domains of risk (e.g., the health domain) contain greater ambiguity about the possible outcomes and probabilities than others.

Conclusion

We demonstrate that self-reported likelihoods of engaging in a risky activity reflect a person's actual choice rather well—at least under conditions of clearly defined activities and when collecting both measures in the same session. However, we also found that the coupling between likelihood ratings and actual choice behavior is loosened when part of the characteristics of the choice options are unknown and in older age. We may know our inner daredevil, but not in every guise.

References

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of personality and Social Psychology*, *65*, 272–281.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293–1295.

Becker, A., Deckers, T., Dohmen, T. J., Falk, A., & Kosse, F. (2012). The relationship between economic preferences and psychological personality measures (April 19, 2012). CESifo Working Paper Series No. 3785. Available at SSRN: http://ssrn.com/abstract=2042458

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*, 351–368.

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

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72*, 691–695.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence–accuracy relation in recognition memory. *Psychonomic Bulletin and Review*, *7*, 26–48.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*, 325–370.

Denburg, N. L., Tranel, D., & Bechara, A. (2005). The ability to decide advantageously declines prematurely in some normal older adults. *Neuropsychologia*, *43*, 1099–1106.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review, 100,* 1238–1260.

Downs, D. S., & Hausenblas, H. A. (2005). Elicitation studies and the theory of planned behavior: A systematic review of exercise beliefs. *Psychology of Sport and Exercise, 6,* 1–31.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75,* 643–669. doi:10.2307/1884324.

Estes, W. K., & Maddox, W. T. (1995). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1075–1095.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 709–730.

Frey R., Mata R., & Hertwig R. (2015). The role of cognitive abilities in decisions from experience: Age differences emerge as a function of choice set size. *Cognition, 142,* 60–80.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science, 15,* 346 –350.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, *123*, 21–32.

Goldstein,W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, *94*, 236–254.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, John Wiley and Sons.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology, 24,* 411–435.

Hagger, M. S., Chatzisarantis, N. L., & Biddle, S. J. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: Predictive validity and the contribution of additional variables. *Journal of Sport and Exercise Psychology, 24,* 3–32.

Hanoch, Y., Johnson, J. G., & Wilke, A. (1996). Domain specificity in experimental measures and participant recruitment. *Psychological Science, 17,* 300–304.

Hansson, P., Juslin, P., & Winman, A. (2008). The role of short-term memory capacity and task experience for overconfidence in judgment under uncertainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1027–1042.

Hansson, P., Rönnlund, M., Juslin, P., & Nilsson, L. G. (2008). Adult age differences in the realism of confidence judgments: Overconfidence, format dependence, and cognitive predictors. *Psychology and Aging, 23,* 531–544.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science, 26,* 433–443.

Henninger, D. E., Madden, D. J., & Huettel, S. A. (2010). Processing speed and memory mediate age-related differences in decision making. *Psychology and Aging, 25,* 262–270.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review, 92,* 1644–1655.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science, 310,* 1680–1683.

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review, 107,* 384–396.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126,* 349–370.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human learning and memory, 6,* 107–118.

Ladouceur, R., Jacques, C., Giroux, I., Ferland, F., & Leblond, J. (2000). Analysis of a casino's self-exclusion program. *Journal of Gambling Studies, 20,* 301–307.

Ladouceur, R., Sylvain, C., & Gosselin, P. (2007). Self-exclusion program: A longitudinal evaluation study. *Journal of Gambling Studies, 23,* 85-94.

Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Personality and Individual Differences, 31,* 215–226.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., et al. (2002). Evaluation of a behavioral measure of risk-taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8,* 75–84.

Lichtenstein, S., & Slovic, P. (Eds.) (2006). *The Construction of Preference.* New York: Cambridge University Press.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide.* New York: Lawrence Erlbaum Associates.

Markiewicz, L., & Weber, E. U. (2013). DOSPERT's gambling risk-taking scale predicts excessive stock trading. *Journal of Behavioral Finance, 14,* 65–78.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 851–861.

MacCrimmon, K. R., Wehrung, D. A. (1990). Characteristics of risk taking executives. *Management Science, 36,* 422–435.

Morewedge, C. K., Gilbert, D. T., Keysar, B., Berkovits, M. J., & Wilson, T. D. (2007). Mispredicting the hedonic benefits of segregated gains. *Journal of Experimental Psychology: General, 136,* 700–709.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology, 15,* 263–280.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5,* 411–419.

Pachur, T., Hanoch, Y., & Gummerum, M. (2010). Prospects behind bars: Analyzing decisions under risk in a prison population. *Psychonomic Bulletin and Review, 17,* 630-636.

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General, 143,* 2000-2019.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137,* 615-625.

Rhodes, M. G., & Jacoby, L. L. (2007). On the dynamic nature of response criterion in recognition memory: Effects of base rate, awareness, and feedback. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 305–320.

Rolison, J. J., Hanoch, Y., & Gummerum, M. (2013). When opportunity matters: Comparing the risk-taking attitudes of prisoners and recently released ex-prisoners. *Risk Analysis, 33,* 2013-2022.

Rolison, J. J., Hanoch, Y., & Wood, S. (2012). Risky decision making in younger and older adults: The role of learning. *Psychology and Aging, 27,* 129–140.

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.

Rolison, J. J., Morsanyi, K., & O'Connor, P. A. (2015). Can I count on getting better? Association between math anxiety and poorer understanding of medical risk reductions. *Medical Decision Making.*

Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

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

Wichary, S., Pachur, T., & Li, M. (2015). Risk-taking tendencies in prisoners and nonprisoners: Does gender matter? *Journal of Behavioral Decision Making, 28,* 504–514.

Zamarian, L., Sinz, H., Bonatti, E., Gamboz, N., & Delazer, M. (2008). Normal aging affects decisions under ambiguity, but not decisions under risk. *Neuropsychology, 22,* 645–657.

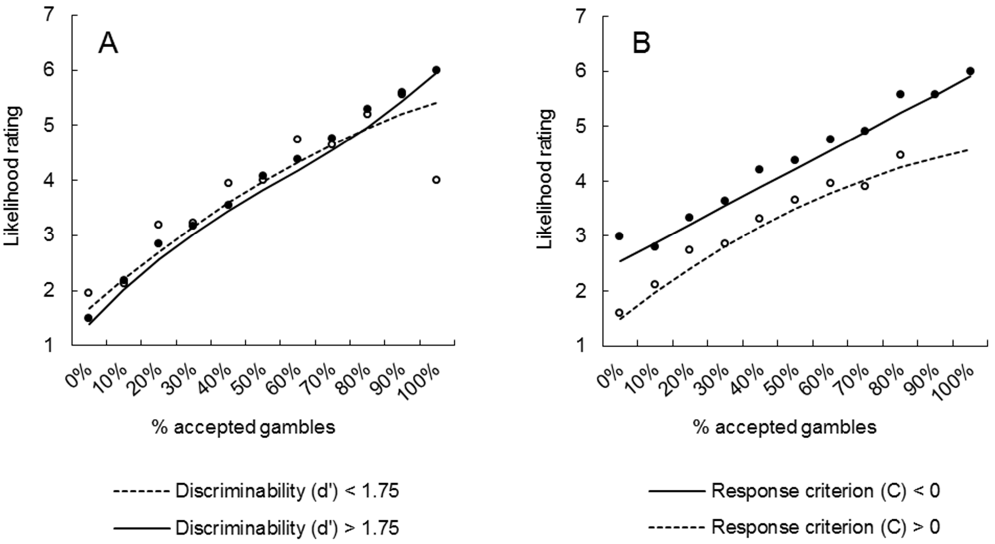

Figure 1. The dots represent the average likelihood rating as a function of the (binned) percentage of accepted gambles at (A) low and high levels of discriminability and (B) response criterion values above and below zero. The lines represented the predicted slopes represent quadratic and cubic effects of accepted gambles on likelihood ratings, respectively.

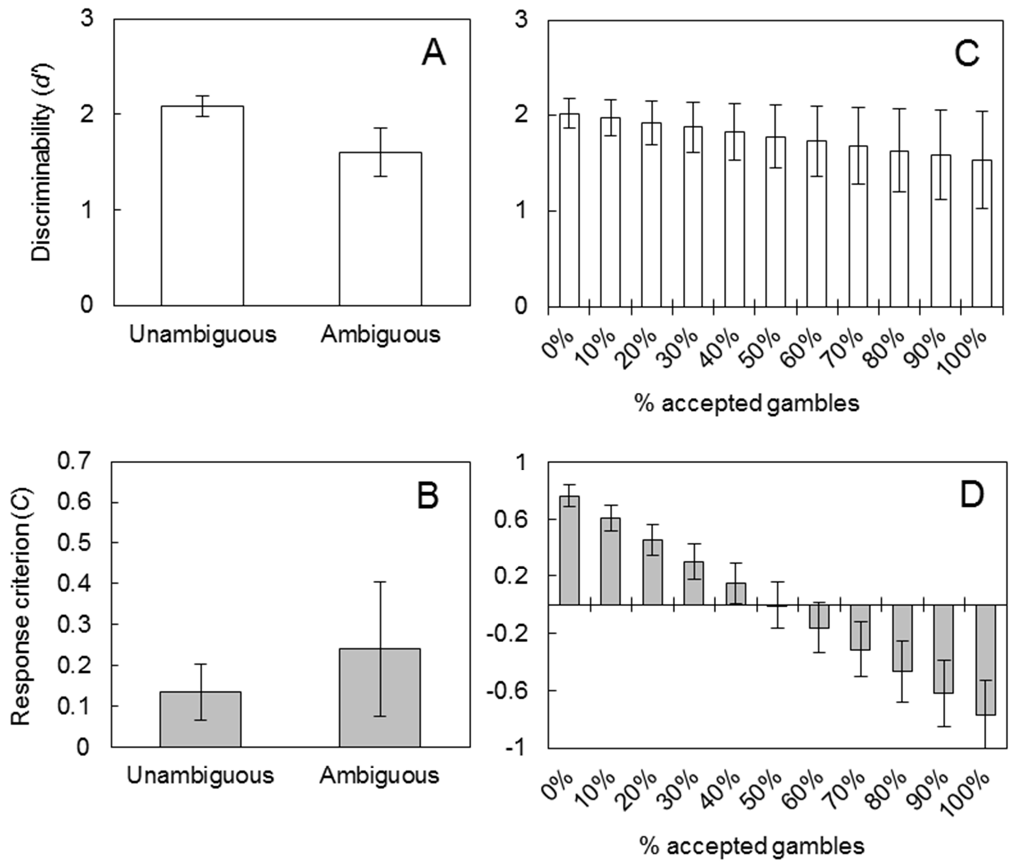

Figure 2. Predicted discriminability and response criterion for (A and B) the unambiguous and ambiguous conditions and (C and D) individual differences in risk taking, measured as the percentage of accepted gambles. Vertical bars indicate the 95% confidence intervals.

Appendix A

Table A1. Unambiguous and ambiguous gambles

Unambiguous gambles					Ambiguous gambles				
Gain	Loss	Win	Lose	Expected value	Gain	Loss	Win	Lose	Expected value
10	10	25%	75%	-5	?	10	75%	25%	12.5
10	20	25%	75%	-12.5	?	10	50%	50%	5
10	30	25%	75%	-20	?	10	25%	75%	-2.5
20	10	25%	75%	-2.5	?	20	75%	25%	10
20	20	25%	75%	-10	?	20	50%	50%	0
20	30	25%	75%	-17.5	?	20	25%	75%	-10
30	10	25%	75%	0	?	30	75%	25%	7.5
30	20	25%	75%	-7.5	?	30	50%	50%	-5
30	30	25%	75%	-15	?	30	25%	75%	-17.5
10	10	50%	50%	0	30	?	75%	25%	17.5
10	20	50%	50%	-5	30	?	50%	50%	5
10	30	50%	50%	-10	30	?	25%	75%	-7.5
20	10	50%	50%	5	20	?	75%	25%	10
20	20	50%	50%	0	20	?	50%	50%	0
20	30	50%	50%	-5	20	?	25%	75%	-10
30	10	50%	50%	10	10	?	75%	25%	2.5
30	20	50%	50%	5	10	?	50%	50%	-5
30	30	50%	50%	0	10	?	25%	75%	-12.5
10	10	75%	25%	5	30	10	?	?	10
10	20	75%	25%	2.5	30	20	?	?	5
10	30	75%	25%	0	30	30	?	?	0
20	10	75%	25%	12.5	20	10	?	?	5
20	20	75%	25%	10	20	20	?	?	0
20	30	75%	25%	7.5	20	30	?	?	-5
30	10	75%	25%	20	10	10	?	?	0
30	20	75%	25%	17.5	10	20	?	?	-5
30	30	75%	25%	15	10	30	?	?	-10

Appendix B

Instructions used in for unambiguous gambling problems:

Thank you for agreeing to take part in our study. The study explores how people think about uncertain outcomes.

We have designed a set of gambles that we would like you to evaluate. Each gamble has two possible outcomes (a win or a loss). Each outcome is characterized by an amount ($10, $20, or $30) that can be won or lost and a chance (i.e., probability) of winning or losing (25%, 50%, or 75%):

(a) win or loss amount ($10, $20, $30)
(b) chance of winning or losing (25%, 50%, 75%)

Here is an example of the kind of gamble you will be shown:

Gamble:	You win	$10 with a chance of 25%
	You lose	$30 with a chance of 75%

To help you understand these chances, you can think of a bag containing 100 tokens, of which 25 are blue and the remaining 75 are red. Imagine drawing one of the tokens from the bag without looking. If you draw one of the 25 blue tokens you win $10. If you draw one of the 75 red tokens you lose $30.

In total, you will be shown 54 such gambles, divided into two blocks. For one block, you will be asked whether or not you would accept each gamble. For another block, you will instead be asked how likely you would be to accept each gamble. You may begin with either block. Finally, you will be asked 5 short demographic questions.

Instructions used in for ambiguous gambling problems:

Thank you for agreeing to take part in our study. The study explores how people think about uncertain outcomes.

We have designed a set of gambles that we would like you to evaluate. Each gamble has two possible outcomes (a win or a loss) that occur with some probability. The outcome can be one of three amounts of money, either $10, $20, or $30, that can be either won or lost. The chance (i.e., probability) of winning or losing can be either 25%, 50%, or 75%.

To help you understand these chances, you can think of a bag containing 100 tokens. When, for instance, the chance of winning is 75% and the chance of losing is 25%, there are 75 blue tokens and the remaining 25 are red. Imagine drawing one of the tokens from the bag without looking. If you draw one of the 75 blue tokens you win the specified amount. If you draw one of the 25 red tokens you lose the specified amount.

For each gamble, either the gain amount ($10, $20, $30), loss amount ($10, $20, $30), or the chance of winning or losing (25%, 50%, 75%) will be unknown.

Here is an example of the kind of gamble you will be shown:

Example 1:

Gamble: You win $10 with a chance of 25%

You lose $? with a chance of 75%

In this gamble, you have a 25% chance of winning $10 and a 75% chance of losing an unknown amount of either $10, $20, or $30.

Example 2:

Gamble: You win $20 with a chance of ? %

You lose $10 with a chance of ? %

In this gamble, you have a chance of winning $20 or to lose $10, but the probability of winning or losing is unknown.

You will first be shown 54 such gambles, divided into two blocks. For one block, you will be asked whether or not you would accept each gamble. For another block, you will instead be asked how likely you would be to accept each gamble. You may begin with either block. You will then be asked to evaluate a final set of 27 gambles. Finally, you will be asked 4 short demographic questions.

Appendix C

We conducted a mixed-effects logistic regression on participants' decisions (accept vs. reject) and included the gamble's expected value, the condition (ambiguous vs unambiguous), and participants' age (as a continuous variable) as predictors. Information about either the gain amount, loss amount, or the chances to win and lose was missing on ambiguous gambles. However, participants were told that the missing gain or loss amount was equal to \$10, \$20, or \$30, and that the missing chances were 25%, 50%, or 75%. Thus, we calculated the expected value of ambiguous gambles by substituting the missing information with the middle amount (i.e., \$20) and probability (i.e., 50%).[1] Gambles with a higher expected value were more often accepted (b = 0.23, t = 51.45, p < .001; Panel A in Figure A1). Ambiguous gambles were less often accepted (34%) than unambiguous (40%) gambles (b = -0.44, t = 2.29, p = .022; Panel A in Figure A1), indicating ambiguity aversion. As age increased, fewer gambles were accepted (b = -0.02, t = 2.51, p = .012). Two-way interaction terms were included in a second block and revealed an interaction between condition and the expected value of the gambles (b = -0.07, t = 7.80, p < .001). This is because participants were less responsive to changes in the expected value of ambiguous gambles (b = 0.20, t = 33.95, p < .001) than they were for unambiguous gambles (b = 0.28, t = 36.95, p < .001). Panel A in Figure A1 shows that this was true particularly when the expected value was positive, further indicating that participants' ambiguity aversion was partly driven by their pessimistic beliefs about the missing information. Age also interacted with the expected value of the gambles (b = -0.001, t = 2.11, p = .035), whereby older age was associated with a reduced sensitivity to a gamble's expected value.

[1] The expected values of ambiguous gambles provided a better fit in the regression model when based on the middle amounts (\$20) and chances (50%) than when based on participants' judgments about the most likely missing values. Nonetheless, participants' mean judgments for the missing values reflected their risk aversion for ambiguous gambles. They judged a missing gain as equally likely to be small (32%), medium (36%), or large (33%), but judged a missing loss as more likely to be medium (36%) or large (40%) than small (24%), and judged a missing chance to win as more likely to be low (36%) or medium (40%) than high (24%) in probability.

In addition to deciding whether to accept or reject the gambles, participants rated in a separate block the likelihood that they would accept each one. We conducted a mixed effects linear regression on their likelihood ratings and included the gamble's expected value, the condition (ambiguous vs unambiguous), and participants' age (as a continuous variable) as predictors. In keeping with our analysis of participants' decisions, the expected values of ambiguous gambles were calculated by substituting the missing information with the middle amount ($20) and probability (50%) on the scale of possible values. Participants rated a higher likelihood that they would accept lotteries with a higher expected value ($b = 0.12$, $t = 92.90$, $p < .001$; Panel B in Figure A1). As age increased, participants rated a lower likelihood of accepting lotteries ($b = -0.01$, $t = 2.23$, $p = .026$). Overall, participants rated that they were less likely to accept ambiguous lotteries ($M = 3.19$, $SD = 1.97$) than unambiguous lotteries ($M = 3.42$, $SD = 2.19$), but this difference was not significant ($b = -0.18$, $t = 1.80$, $p = .073$; Panel B in Figure A1). However, when two-way interaction terms were included in a second block, condition interacted with the expected value of the lotteries ($b = -0.02$, $t = 8.65$, $p < .001$). This was because participants were less responsive to changes in the expected value of ambiguous lotteries ($b = 0.02$, $t = 46.03$, $p < .001$) than they were for unambiguous lotteries ($b = 0.13$, $t = 75.58$, $p < .001$; Panel B in Figure A1). Finally, there was a marginally significant interaction between age and the expected value of the lotteries ($b = -0.0002$, $t = 1.91$, $p = .056$), indicating that older age was associated with a reduced sensitivity to a gamble's expected value.

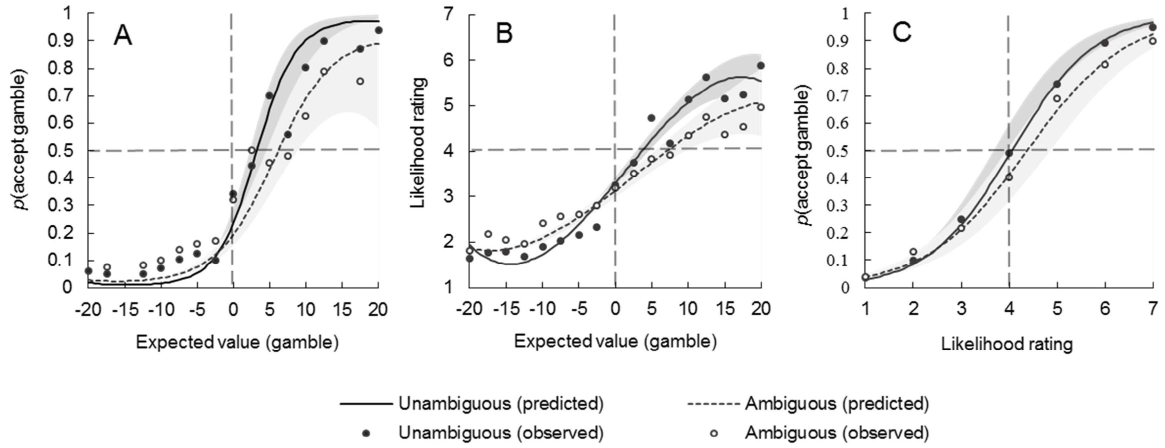

Figure A1. Predicted probability (A) and rated likelihood (B) of accepting gambles according to their expected value and predicted probability (C) of accepting gambles according to likelihood ratings. The predicted slopes represent cubic effects of expected value on the probability of accepting gambles and likelihood ratings. The dots represent mean group values.