

Review



Cite this article: Hertwig R, Wulff DU, Mata R. 2019 Three gaps and what they may mean for risk preference. *Phil. Trans. R. Soc. B* **374**: 20180140.
<http://dx.doi.org/10.1098/rstb.2018.0140>

Accepted: 4 October 2018

One contribution of 14 to a theme issue ‘Risk taking and impulsive behaviour: fundamental discoveries, theoretical perspectives and clinical implications’.

Subject Areas:

behaviour, cognition

Keywords:

risk preference, description–experience gap, revealed preferences, stated preference, convergent validity, temporal stability

Author for correspondence:

Ralph Hertwig
e-mail: hertwig@mpib-berlin.mpg.de

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4305470>.

Three gaps and what they may mean for risk preference

Ralph Hertwig¹, Dirk U. Wulff^{1,2} and Rui Mata^{1,2}

¹Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

²Department of Psychology, University of Basel, 4055 Basel, Switzerland

RH, 0000-0002-9908-9556; DUW, 0000-0002-4008-8022; RM, 0000-0002-1679-906X

Risk preference is one of the most important building blocks of choice theories in the behavioural sciences. In economics, it is often conceptualized as preferences concerning the variance of monetary payoffs, whereas in psychology, risk preference is often thought to capture the propensity to engage in behaviour with the potential for loss or harm. Both concepts are associated with distinct measurement traditions: economics has traditionally relied on behavioural measures, while psychology has often relied on self-reports. We review three important gaps that have emerged from work stemming from these two measurement traditions: first, a description–experience gap which suggests that behavioural measures do not speak with one voice and can give very different views on an individual’s appetite for risk; second, a behaviour–self-report gap which suggests that different self-report measures, but not behavioural measures, show a high degree of convergent validity; and, third, a temporal stability gap which suggests that self-reports, but not behavioural measures, show considerable temporal stability across periods of years. Risk preference, when measured through self-reports—but not behavioural tests—appears as a moderately stable psychological trait with both general and domain-specific components. We argue that future work needs to address the gaps that have emerged from the two measurement traditions and test their differential predictive validity for important economic, health and well-being outcomes.

This article is part of the theme issue ‘Risk taking and impulsive behaviour: fundamental discoveries, theoretical perspectives and clinical implications’.

1. Introduction

The construct of risk preference is one of the most important building blocks of economic and psychological theories of choice. It is often invoked to explain behaviours and interindividual differences therein in domains as diverse as individuals’ financial choices, unlawful behaviours (e.g. speeding, tax evasion), health choices (e.g. consuming recreational and possibly illicit drugs) and professional choices (e.g. entrepreneurial initiatives). Risk preference—also termed ‘risk attitude,’ ‘risk tolerance’ or ‘sensitivity to risk’—is often understood to represent a personal characteristic. Despite this default view, opinions about how to best conceptualize this construct vary [1,2], including whether it represents a stable individual characteristic [3,4], how it relates to other mainstay psychological constructs such as impulsivity [5,6] and how individual differences in risk preference should be measured [1,7,8].

In what follows, we distinguish between two major measurement traditions of investigating individual differences in risk preference. One originates in economics and rests on behavioural measures, such as the choice between monetary gambles. Another, originating in psychology, tends to rely on self-reports. As we argue below, understanding the two approaches and their somewhat conflicted relationship is key to assessing the current literature on risk preference. We then review evidence which suggests that the two sets of measures

provide different and almost opposing views about the nature of risk preference. Specifically, we introduce three gaps between measurements of risk preferences in economics and psychology that have been identified recently: first, a gap in the behavioural patterns observed for monetary gambles presented in different formats, the *description–experience gap*; second, a gap in the convergent validity of behavioural and self-report measures, the *behaviour–self-report gap*; and, third, a gap in the observed temporal stability of behavioural and self-report measures, the *temporal stability gap*. We conclude that future work must reconcile the contradictory findings produced by the two measurement traditions. In particular, we propose that a systematic and evidence-based understanding of the relation between different measures will be crucial for making conceptual and empirical progress in the study of risk preferences.

2. What is risk preference?

When economists and psychologists call behaviours ‘risky’ they use the same term but mean different things. In economics and finance, risk preference commonly refers to the tendency to choose an action that involves higher variance in potential monetary outcomes, relative to another option with a lower variance of outcomes (but equal expected value). This holds independent of whether these outcomes involve gains or losses [9,10]. For example, when offered the choice between a safe option of receiving €500 guaranteed and a risky option of a 50% chance of receiving €1000 and a 50% chance of receiving nothing, a risk-neutral person would not prefer one option over the other. Expected value maximization (i.e. multiplying all outcomes per options with their respective probabilities, summing the products and maximizing), which embodies risk neutrality, values both options equally. However, other modelling frameworks make it possible to consider risk-averse and risk-loving individuals as well. For example, according to expected utility theory, a classic model in economics [11], a risk-averse person may be said to possess a concave utility function that leads to a preference for choosing the safe option, implying that gaining €500 contributes more than half the utility of a 50% chance of gaining €1000, thus accommodating their choice of the sure option. Likewise, a risk-prone person’s utility curve may be said to be convex. There is a long tradition of using such mathematical theories in economics [12], with several competing formulations now available to capture individuals’ risk preferences [13]. Let us highlight and clarify that utility-based economic modelling of human behaviour has been applied not only to risky activities such as financial investments but also to ‘choices’ such as drug use (addiction [14]) and criminal behaviours: as Becker [15] put it: ‘a person commits an offense if the utility to him exceeds the utility he could get by using his time and other resources at other activities’ (p. 176).

In psychology, risk preference is often broadly interpreted as the propensity to engage in behaviours or activities that, although rewarding, involve the potential for loss or harm (for oneself or others). Psychologists have shown less interest than have economists in comparing the implications of different mathematical formulations of utility for risk preference (with prominent exceptions; [16]), focusing instead on understanding whether latent constructs derived from self-reports

of attitudes and behaviours are associated with drug use or daring activities such as speeding, rock climbing and imprudent online behaviour—all of which may be rewarding but also carry the possibility of physical or psychological harm (e.g. [17]).

Although we adopt the umbrella term ‘risk preference,’ we note that a number of distinct but overlapping psychological constructs have been used to account for individual differences in such ‘risky’ behaviours, including impulsivity [18], sensation-seeking [19], novelty-seeking [20] and impulse control [17]. Further, a number of psychological theories concerning these constructs do not necessarily propose them as unitary. For example, the tripartite view on impulsivity suggests that it consists of components of reward sensitivity, loss sensitivity and inhibitory control (see [21]). Ultimately, the degree of overlap between these constructs will be largely an empirical question—an issue we discuss in more detail below.

3. Two measurement traditions: revealed and stated preferences

One would expect that a construct as important and frequently invoked as risk preference rests on a firm measurement foundation, but this is far from the truth. Indeed, the conceptualizations of risk preference in economics and psychology gave rise to two distinct ways of measuring risk preference (for reviews, see [8,22,23]).¹ How these measures relate to each other has, at least until recently, received scant attention. The *revealed-preference tradition* [26,27] holds that people’s utilities and true beliefs are revealed through the (incentivized) choices they make.² Consequently, this tradition has relied predominantly on simple monetary gambles (e.g. [16,28,29]) or extensions of it, such as the multiple-price-list method [30]. It is unsurprising that choices among monetary gambles have played an outsized role in measuring risk preference in economics: they were midwives of the Enlightenment concept of mathematical expectation [31]; they gave rise, through the St. Petersburg gamble, to what is known today as expected utility theory [11]; they were invoked to demonstrate that people’s choices are at odds with axioms of expected utility theory (e.g. [28,32]); they were enlisted to demonstrate the key concepts of prospect theory such as loss aversion and the fourfold pattern of risk preference [16]; and last but not least, one simple gamble type—the choice between a safe and a risky option—has frequently been employed to gauge people’s risk preference.

Proponents of the *stated-preference tradition*, by contrast, bank on people’s introspective abilities rather than their observable behaviour. They elicit data concerning risk using relatively general questions (‘Are you generally a risk-taking person or do you try to avoid risks?’ [33]), specific but hypothetical questions (‘How likely would you be to go whitewater rafting at high water in the spring?’ [34]), or ask people to report on the frequency of actual risky activities (‘How many cigarettes do you smoke per day?’). Self-report measures have been widely used in applied and epidemiological contexts, presumably because they are easy to administer. For example, financial institutions often rely on self-report measures to gauge their clients’ risk preference so as to meet the legal requirements for the sale of financial

products [35]. Self-report measures of risk preference have commonalities with those used to measure overlapping constructs such as impulsivity [21] and sensation-seeking [36].

Is the existence of these two measurement traditions consistent with a single latent trait that comes into sight regardless of how it is being probed? Or do the measures ‘make’ the construct—do two distinct constructs surface if these measures are employed simultaneously? Recent years have seen considerable efforts to assess the operation and implications of these measures, including their convergent validity [5]; their temporal stability [5,8]; their associations with personality traits and demographic characteristics such as age, gender and cognitive abilities [37,38]; and their genetic basis [39,40]. In what follows, we describe three major findings involving measures stemming from both traditions that each suggest an incongruity to be addressed.

4. The description–experience gap

Monetary gambles are first among equals in behavioural measures. Interestingly, there is little variability in their guise. Researchers typically present the options’ outcomes and probabilities numerically (e.g. €500 guaranteed versus €1000 with 0.5; €0 with 0.5) or with a spinner wheel or bar chart (see the meta-analysis from [41]), and gambles explicitly state all possible outcomes and their probabilities. This nearly invariant choice architecture for measuring people’s response to risk is rather odd because ‘it is hard to think of an important natural decision for which probabilities are objectively known’ ([42], p. 325). Indeed, in everyday life, people rarely encounter convenient descriptions of objective probability distributions, with a few exceptions such as the probability of rain (e.g. [43]). Instead, people must turn to whatever experience they may have, making *decisions from experience* rather than the *decisions from description* that are often studied in laboratories [44,45].

Decisions from description versus decisions from experience—this simple distinction, which should be understood more as poles on a continuum rather than a dichotomy, raises a new question: do these two modes of learning about the probabilistic texture of the world [46] result in the same or systematically different choices? The question has received much attention since three articles in the early 2000s [41,44,47] demonstrated a systematic discrepancy in description- and experience-based choices: the *description–experience gap* (for reviews, see [48–50]). These and many subsequent studies presented both gambles in which all outcomes and probabilities are stated (description) and gambles in which the payoff distributions were initially unknown but people could randomly draw from them (experience). Each draw produced one outcome; draw-by-draw, the properties of the outcome distributions were revealed. Two major experiential paradigms have been employed (although many hybrid variants exist): in the *sampling paradigm*, people first sample as many outcomes as they wish, then decide from which distribution to make a single draw. In the *partial-feedback paradigm*, each of a typically large, fixed number of draws contributes to people’s earnings and they receive draw-by-draw feedback on the obtained payoffs. The sampling paradigm removes the exploitation–exploration tradeoff [51], whereas the partial-feedback paradigm incorporates it. Assuming that people sample sufficiently and

equally across the payoff distribution, the description and experience offer equivalent information. But are the resulting choices equivalent?

(a) Manifestations of the description–experience gap

As it turns out, they are often not equivalent. There are several ways to illustrate the description–experience gap [49,50]. Figure 1 plots the gap in terms of a systematic difference in the observed choice proportions in description and experience, as a function of gamble type (and for a subset of studies that examined the gap rather than its boundary conditions; see [45] for how the systematic differences were determined).³ When a choice involves a risky and a safe option—the choice task often used to behaviourally measure risk preference—the gap is 18.7 percentage points; when a choice involves two risky options the gap is 7 percentage points.

A second manifestation of the description–experience gap pertains to the maximization rate. In decisions from description, Wulff *et al.* [45] found a median of 55% of choices maximized expected value; in decisions from experience, by contrast, 66% of people who encountered all possible outcomes and 89% of people who experienced some, but not all outcomes (often missing the rare event) maximized the experienced mean return—that is, the ‘expected value’ of the actually experienced sample of outcomes. A third way to demonstrate that choices in description and experience are systematically different—and one that brings us straight back to risk preference—focuses on the *fourfold pattern of risk preference* [52]. This pattern is shown in table 1. The classic model of decisions under risk, expected utility theory, assumes that individuals are generally risk-averse (i.e. a concave utility function). Challenging this view, Tversky & Kahneman [52] showed that people are both risk-averse and risk-seeking. Consider the choice proportions in gambles with stated probabilities. In the gain domain, most people were risk-averse, preferring the safe option when the probability of winning was high (4 with 0.8). When the gamble had the same expected value but a low probability of winning (32 with 0.1), preference reversed: most people were risk-seeking and chose the risky option. With the same choices but with outcomes in the loss domain, preferences flipped again: many people proved risk-averse when the stated probability of losing was low (–32 with 0.1) but risk-seeking when it was high (–4 with 0.8). Now consider what happens when people make decisions from experience with the same options: the fourfold pattern reverses, suggesting that in decisions from description, people choose *as if* they tend to overweight rare events, whereas in decisions from experience they choose *as if* they underweight rare events; we return to the implied probability weighting shortly.

(b) What contributes to the description–experience gap?

The evidence reviewed above suggests there is a description–experience gap in choice proportions, but what causes it? Several determinants have been examined [48]; here we briefly review two of the most extensively studied explanations. The first is reliance on small samples. Indeed, Wulff *et al.* [45] found that across many thousands of trials, the median sample size was 14 across trials with one safe and one risky option, and 22 across trials with two risky

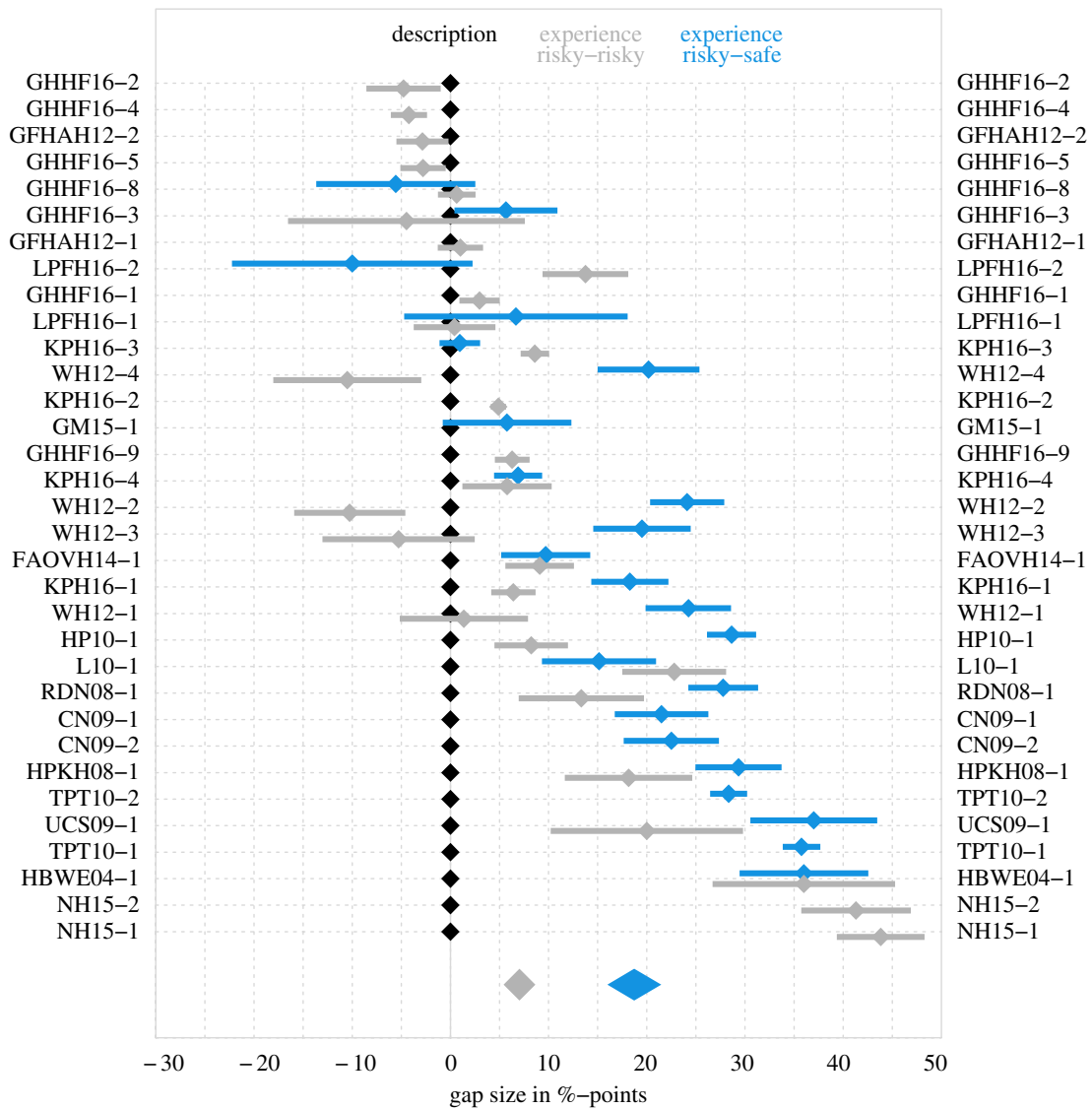


Figure 1. The magnitude of the description–experience gap. The difference in the proportions of choices consistent with discrete underweighting in decisions from experience minus those in decisions from description, from a variety of published and unpublished studies. Results are based on the data compiled in Wulff *et al.* [45] and are restricted to studies that aimed to measure the gap rather than study its boundary conditions. The results are shown for choice problems involving a risky and safe option (darker) versus two risky options (lighter). The references behind the acronyms can be found in Wulff *et al.* [45] or received from the authors of this paper. Adapted from [45]. (Online version in colour.)

Table 1. The fourfold pattern of risk attitudes in decisions from description and its reversal in decisions from experience (based on [44,53]).

probability	description		experience	
	gain domain	loss domain	gain domain	loss domain
low	32, 0.1 ^a versus 3, 1.0 rare event: 32, 0.1 risk-seeking 48% ^b	–32, 0.1 versus –3, 1.0 rare event: –32, 0.1 risk-averse 36%	32, 0.1 versus 3, 1.0 rare event: 32, 0.1 risk-averse 20%	–32, 0.1 versus –3, 1.0 rare event: –32, 0.1 risk-seeking 72%
high	4, 0.8 versus 3, 1.0 rare event: 0, 0.2 risk-averse 36%	–4, 0.8 versus –3, 1.0 rare event: 0, 0.2 risk-seeking 72%	4, 0.8 versus 3, 1.0 rare event: 0, 0.2 risk-seeking 88%	–4, 0.8 versus –3, 1.0 rare event: 0, 0.2 risk-averse 44%

^aThe alternative outcome (0 otherwise) has been omitted for all risky options.

^bThis is the proportion of risky choices observed. In past studies, this proportion (48%) has been found to be greater than 50% (e.g. [52]).

options. Relatively modest sampling effort exacts a price. In about one third of trials, people did not experience at least one of the possible outcomes—typically the rare event. Consequently, people relied, on average, on small samples that caused systematically distorted representations of the true probabilities. Yet sampling error is not the sole determinant of the gap. In an analysis of trials in which the experienced frequencies closely tracked the true probabilities, Wulff *et al.* ([45], fig. 7) nevertheless observed a description–experience gap. Sampling error may thus be sufficient but not necessary for the gap to emerge (see also [54]).

A second factor that has received much attention is the weighting of stated and experienced probabilities (i.e. relative frequencies). In a recent sophisticated analysis of the weighting of the *objective* probability of choice options, Regenwetter & Robinson [55] found strong evidence for a gap: people overweighted rare events in choices from description and underweighted rare events in those based on experience (consistent with early conclusions about the gap; see [44,47]). In experience, however, people do not have access to options' objective probabilities. Therefore, many studies analysed the weighting of the actually experienced probabilities. In an exploratory analysis of the meta-analytical dataset, Wulff *et al.* ([45]; fig. 10) found that experienced relative frequencies and stated probabilities prompted different weighting functions—more linear weighting in experience versus overweighting of rare events in description—for choices involving a risky and a safe option, but similar overweighting for choices with two risky options.

Going beyond search (i.e. small samples) and probability weighting explanations, researchers concerned with the description–experience gap have proposed several new models to account for experiential choice (e.g. [56]; for a review, see [48]). Together these models suggest that description- and experience-based choices engage different cognitive processes: unlike in description, experience-based choices are based on sequential search and updating and may involve rudimentary or no explicit use of probabilities [45,57]. Figure 2 offers another illustration of the diverging processes in description and experience. Using the meta-analytical data compiled in Wulff *et al.* [45], we analysed the extent to which 12 different models of choice predict people's choices in description and experience. Two results are noteworthy: first, in decisions from experience, a simple heuristic—the natural-mean heuristic (see electronic supplementary material)—performs as well as or even better than cumulative prospect theory; a Bayesian model; and two mechanistic explanations specific to decisions from experience, round-wise integration [58,59] and the ExCon model [60]. The natural-mean heuristic reaches the same choice as expected value theory (applied to the 'experienced' data) would, but it does so without any multiplication or explicit representation of probabilities. In decisions from description, by contrast, this heuristic falls far behind. Here, however, another simple heuristic, which weighs all distinct outcomes per gamble equally (the equiprobable heuristic; electronic supplementary material), performs nearly as well as cumulative prospect theory [52].

To conclude, people's choices and revealed risk preferences are systematically different in description and experience. The description–experience gap has an important implication for risk preference: depending on how individuals learn about their options they may employ different

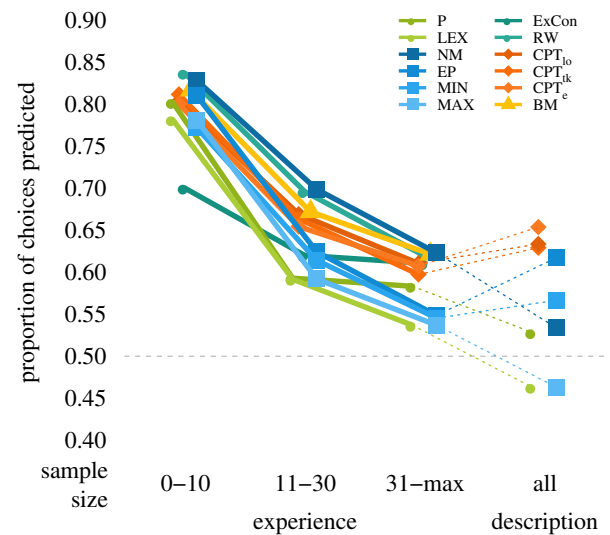


Figure 2. Predicting decisions from experience and description. The plot shows the predictive accuracy of 12 choice models tested using choice problems with a risky and a safe option. In decisions from experience, the choice proportion is calculated as a function of three categories of sample sizes (0–10, 11–30, greater than 30). Data include 15 054 experience- and 10 239 description-based choices compiled in Wulff *et al.*'s [45] meta-analysis. The model competition includes six heuristics (NM, P, LEX, EP, MIN, MAX), three variants of cumulative prospect theory (CPT_{lo} , CPT_{th} , CPT_e), a Bayesian updating model (BM), round-wise integration (RW) and the exemplar confusion model (ExCon). Details of the models are described in the electronic supplementary material.

cognitive processes, thereby arriving at different decisions. In this regard, the description–experience gap bears similarity to other, classic format effects such as gain–loss framing [61] or branch-splitting [32,62], except that the description–experience gap has been found to generalize beyond monetary gambles to, for instance, causal reasoning (e.g. [63]), consumer choice [59] or Bayesian reasoning [64]. The description–experience gap has also been observed in nonhuman primates, thus suggesting that it ‘does not depend on uniquely human cognitive abilities, such as those associated with language,’ and supporting the idea that ‘epistemic influences on risk attitudes are evolutionarily ancient’ ([65], p. 593). In any case, format effects demonstrate how difficult it is to estimate general risk preferences from behavioural choices between monetary gambles, most likely because different formats tap into different decision strategies and associated cognitive mechanisms, as figure 2 also illustrates. Do other tools, such as self-report measures and behavioural tasks other than monetary gambles, produce similar levels of dissonance? We turn to this question next.

5. The behaviour–self-report gap

Psychologists and economists measure risk preference employing either behavioural or self-report measures, rarely relying on both simultaneously. This need not be a problem if the measures converge toward equivalent conclusions about this latent attribute of a person. But do they? Frey *et al.* [5] investigated this crucial question in what is likely the most comprehensive study of risk-preference measures so far, involving 1507 participants who responded to an extensive battery of self-report *and* behavioural measures of risk preference. The authors implemented a psychometric

(bifactor) model to gauge how responses to 39 measures of risk preference are interrelated and estimate how much variance across these could be accounted for by a single general factor of risk preference, *R* (for a complete list of the measures see table 1 in Frey *et al.* [5]). In this model, *R* directly accounts for somewhat more than half of the explained variance across all the measures employed. In addition, self-report measures gauging impulsivity and sensation-seeking also loaded on the *R* factor, highlighting the empirical overlap of operationalizations of risk preference and other such constructs. Frey *et al.*'s [5] analysis is in line with the robust observation that a small set of traits appears to account for large portions of variance in psychological constructs [66,67]. In addition, another set of six domain-specific factors representing risk taking in domains such as health, finances and recreation accounted for the remaining explained variance. Overall, this analysis contrasts with the idea of risk preference as a purely general or domain-specific construct and suggests instead that risk preference encompasses both general and domain-specific components (see also [68]).

Frey *et al.*'s [5] comprehensive psychometric analysis of risk-preference measures produced yet another consequential observation. The general *R* factor did not generalize to the behavioural measures of risk preference and, in fact, did not account for any significant amount of variance in those measures. More broadly, the eight behavioural measures investigated failed to converge (correlate) not only with the self-report measures but also with each other, suggesting that disparate behavioural measures do not capture the same construct of risk preference. We refer to the different patterns of convergent validity between self-report and behavioural measures as the *behaviour–self-report gap*. Convergent validity refers to the degree to which measures of a psychological construct capture a common underlying characteristic or trait—something that risk preference measurement traditions, in particular the behavioural measures, seem to lack. Overall, the *behaviour–self-report gap* raises numerous questions for the future; let us consider just three.

First, what exactly do the behavioural measures capture and what causes the high level of inconsistency between them? One possibility is that the behavioural measures, unlike self-report measures, capture *transient* states rather than a *stable* preference. Their inconsistencies may be a result of them engaging, to different degrees, various decision strategies and cognitive processes such as memory and learning [69], as we discussed in our examination of the description–experience gap. Another possibility is that the ‘preferences revealed are not independent of the procedure (institution) through which they are revealed’ ([70], p. 4213), and this contingency on the elicitation procedure may be more pronounced for revealed-preference measures than for stated-preference measures. Finally, it may be that behavioural measures lack the necessary reliability for consistency between measures to emerge. Recent analyses, for instance, have demonstrated that at least 200 choices between monetary lotteries—many more than are usually employed—are required in order to reliably identify whether individuals weigh probabilities linearly ([71]; see also [72]).

Second, is the extensive empirical overlap between self-reports the product of valid or artefactual (e.g. biased) reporting? Individuals do not have perfect insight into their own personalities and behaviours [73]) and some of the

consistency across self-report ratings of different measures could be associated with recall biases (e.g. the type of events recalled) or response biases (e.g. how rating scales are used; [74]). Overall, the psychological literature suggests that self-reports are mostly valid, albeit imperfect, measures of behaviour [73,75]. However, more work needs to be conducted to assess how some conclusions about the psychometric structure of risk preference [5] can be replicated using other measures, such as informant reports or other, objective measures. For example, field studies suggest some but limited consistency across measures of risk preference estimated from different domains (e.g. car and house insurance; see [1] for an overview); it would be important to document how the convergent validity of such risk preference indices map onto the convergent validity observed for self-report measures.

A third question is whether this striking behaviour–self-report gap is unique to the measurement of risk preference or if it generalizes to other constructs that are also typically gauged via both behavioural and self-report measures such as self-control and impulsivity (e.g. [76]) or social preferences (e.g. [77]). Our reading of the literature suggests that such gaps may be pervasive in the behavioural sciences (e.g. [75,76]) but it would be important to systematically assess how the ontology of risk preference and related constructs differs as a function of measurement choices [78].

6. The temporal stability gap

To the extent that risk preference is considered to be an enduring attribute of a person—tantamount to personality traits [3] or akin to enduring tastes in the classic economic view [2,39]—a pressing issue is whether different measures support the existence of considerable temporal stability of risk preference, or whether they suggest a more capricious attribute that varies substantially with time and resembles more transitory states such as emotions. Systematic variation of these states may, for instance, be a function of the organism's current metabolic needs (as proposed in risk-sensitive foraging theory [79–81]) or aspirations. One measure often used to quantify temporal stability is the test–retest reliability of an individual's risk preferences; this makes it possible to quantify the extent to which the same rank ordering of individuals is preserved across two measurement time points. In a recent meta-analysis of studies reporting test–retest correlations of risk-preference measures, Mata *et al.* [8] found a substantial temporal divergence of behavioural and self-report measures. For choices between monetary gambles, for which no data were available with retest intervals longer than 5 years, correlations of about 0.2 were observed (with considerable variation around this estimate; fig. 1*a* in Mata *et al.* [8]). For self-report measures, by contrast, the corresponding test–retest correlations were around 0.5; in addition, this substantially higher level of stability in the self-report measures appears not to decline further across a 10-year period. These results suggest a third gap between behavioural and self-report measures—this time at the level of the *temporal stability gap*.

Let us emphasize that the considerable stability of risk preference found for self-report measures by no means precludes intraindividual change across the lifespan. Josef *et al.* [38] analysed different notions of stability in a large sample

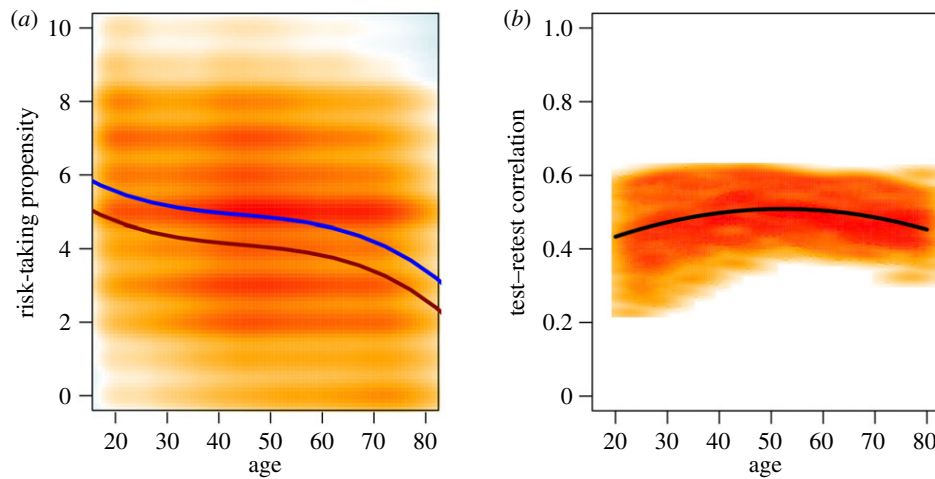


Figure 3. Change and stability in risk preference. The effect of age on (a) cross-sectional mean-level change ('Are you generally a person who is willing to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means not at all willing to take risks and the value 10 means very willing to take risks'), and (b) temporal stability of self-reported risk preference. In (a), the two lines represent the results for males (upper) and females (lower), respectively. All data stem from the German socio-economic panel and are plotted on a kernel density plot in which darker shades indicate a higher density of responses. In (b), the black line represents the weighted average from test–retest correlations between all pairs of measurement waves for males and females combined. Adapted from [38].

of individuals over a period of 10 years, including *differential stability*—consistency in the rank ordering of individuals' risk preferences over time (as in the meta-analysis by Mata *et al.* [8], described above)—and *mean-level stability*—consistency in the respondents' average risk preference over time. Figure 3 plots the findings for these two notions of stability. The results suggest that risk preference—based on a self-report measure—can be understood as a trait with moderate rank-order stability showing correlations of about 0.5 across measures of up to 10 years (figure 3a); however, there may be important lifespan differences in this stability. Specifically, the inverted U-shape pattern in figure 3 suggests that there may be significantly more changes in early adulthood and old age relative to middle age, for example, owing to the shifting nature of social roles and challenges during these phases of life. Similarly, there are reliable mean-level differences across the lifespan, with risk-taking preference typically decreasing across adulthood (figure 3b). Several other results using both cross-sectional and longitudinal analysis support such patterns [37,82].

All in all, the temporal stability gap suggests that self-report and behavioural measures of risk preference show considerable differences in temporal stability, and that self-report measures show important patterns of differential (rank-order) stability and mean-level change that beg investigations of the short- and long-term sources of individual differences in such measures, including biological mechanisms [39,83], cultural origins [84] and the role of specific life events and momentary challenges [4]. We look forward to seeing more work tackling these issues and establishing the convergence or uniqueness of such patterns relative to those for related constructs such as sensation-seeking and impulsivity [17].

One obvious but important point is that differential stability is a prerequisite for convergent validity. Only measures that are reliable across time can be expected to converge and be predictive of other relevant constructs or outcomes of interest. Consequently, the temporal stability gap between self-report and behavioural measures suggests that it could be important to develop more reliable behavioural measures before embarking on projects that compare the predictive power of measure types.

7. Conclusion

To the extent that an individual's risk preference is measurable, one could harness it to, for instance, diagnose and mitigate risk taking that is harmful to the individual and others (e.g. [23,85]). Yet, despite the risk preference construct's time-honoured pedigree [11] and prominent role in economic models of choice [12,13], it is troubled by unresolved and longstanding conceptual and measurement issues. We have reviewed several of these issues and outlined the challenges ahead.

First, focusing on the classic behavioural measure of choice between monetary gambles, we have highlighted the description–experience gap. It suggests that different forms of learning about and representations of the probabilistic options in an environment [46] can engage varying cognitive processes that, in turn, give rise to systematically varied choices and implied risk attitudes [45]. This work is part of a larger stock of findings according to which different presentation formats trigger specific choice regularities, including systematic preference reversals [86]. Second, we have discussed a gap between behavioural and self-report measures of risk preference in terms of their convergent validity. Behavioural measures do not relate to self-report measures and, to add insult to injury, fail to converge with each other; self-report measures, by contrast, show substantial convergent validity. This means that quite different self-report measures, associated with distinct but overlapping constructs (e.g. risk preference, impulsivity and sensation-seeking) and domains (e.g. health, finance and recreation), share large portions of variance. Finally, there is also a gap in the temporal stability of behavioural and self-report measures: only self-report measures show medium to large rank-order stability across years. This is in stark contrast to the argument—typically advanced by researchers who focused primarily on empirical results from monetary gambles (e.g. [86,87])—that risk preference does not represent a stable and trait-like construct.

Although less is known about the predictive validity of risk-preference measures, self-reported measures appear to show significant predictive validity for notable economic

and health outcomes (e.g. teenage pregnancy, drug use; see [24,88]; cf. [8] for an overview), thus dispelling the notion of self-reports being little more than ‘cheap talk.’ At the same time, the picture that emerges from studies with revealed (behavioural) preference measures is less promising (e.g. [89,90]), possibly because they tap into more transient states of risk preference and, furthermore, capture additional cognitive processes such as learning, memory or numeracy skills [8,37]. All in all, it is clear that the field still needs much more work linking different self-report and behavioural measures to field outcomes (cf. [1]).

Theoretical and empirical research on risk preference is more exciting than ever. New questions abound: how and to what degree do the related constructs of, for instance, risk preference and impulsivity, and their measures overlap? Behavioural measures do not appear to be capable of measuring the trait-like characteristics of risk preference, but perhaps they can simulate and predict people’s responses to the specific incentive structure and choice architecture of a real-world context. If so, conceptualizing and testing what is being mustered with each behavioural measure is crucial. And, of course, we need cross-sectional and longitudinal studies that include many of the extant and novel measures in order to quantify their (in)ability to predict important life outcomes such as investment, insurance and health decisions. Risk preference is a central construct shared across the behavioural sciences but current empirical results present a number of puzzling gaps that must be addressed. We propose that it is important to constructively challenge disciplinary preconceptions and measurement traditions to reveal the still enigmatic persona of risk preference.

References

- Barseghyan L, Molinari F, O’Donoghue T, Teitelbaum JC. 2018 Estimating risk preferences in the field. *J. Econ. Lit.* **56**, 501–564. (doi:10.1257/jel.20161148)
- Stigler GJ, Becker GS. 1977 De gustibus non est disputandum. *Am. Econ. Rev.* **67**, 76–90.
- Chuang Y, Schechter L. 2015 Stability of experimental and survey measures of risk, time, and social preferences: a review and some new results. *J. Dev. Econ.* **117**, 151–170. (doi:10.1016/j.jdevco.2015.07.008)
- Schildberg-Hörisch H. 2018 Are risk preferences stable? *J. Econ. Perspect.* **32**, 135–154. (doi:10.1257/jep.32.2.135)
- Frey R, Pedroni A, Mata R, Rieskamp J, Hertwig R. 2017 Risk preference shares the psychometric structure of major psychological traits. *Sci. Adv.* **3**, e1701381. (doi:10.1126/sciadv.1701381)
- Nigg JT. 2016 Attention and impulsivity. In *Developmental psychopathology: Vol. 1. Maladaptation and psychopathology* (ed. D Cicchetti), 3rd edn, pp. 1–56. Hoboken, NJ: Wiley.
- Beshears J, Choi JJ, Laibson D, Madrian BC. 2008 How are preferences revealed? *J. Public Econ.* **92**, 1787–1794. (doi:10.1016/j.jpubeco.2008.04.010)
- Mata R, Frey R, Richter D, Schupp J, Hertwig R. 2018 Risk preference: a view from psychology. *J. Econ. Perspect.* **32**, 155–172. (doi:10.1257/jep.32.2.155)
- Markowitz H. 1952 Portfolio selection. *J. Finance* **7**, 77–91. (doi:10.1111/j.1540-6261.1952.tb01525.x)
- Pratt JW. 1964 Risk aversion in the small and in the large. *Econometrica* **32**, 122–136. (doi:10.2307/1913738)
- Bernoulli D. (1738/1954). Exposition of a new theory on the measurement of risk. *Econometrica* **22**, 23–36. (doi:10.2307/1909829)
- Schoemaker PJH. 1982 The expected utility model: its variants, purposes, evidence and limitations. *J. Econ. Lit.* **20**, 529–563.
- O’Donoghue T, Somerville J. 2018 Modeling risk aversion in economics. *J. Econ. Perspect.* **32**, 91–114. (doi:10.1257/jep.32.2.91)
- Becker GS, Murphy KM. 1988 A theory of rational addiction. *J. Polit. Econ.* **96**, 675–700. (doi:10.1086/261558)
- Becker GS. 1968 Crime and punishment: an economic approach. *J. Polit. Econ.* **76**, 169–217. (doi:10.1086/259394)
- Kahneman D, Tversky A. 1979 Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–292. (doi:10.2307/1914185)
- Steinberg L. 2013 The influence of neuroscience on US Supreme Court decisions about adolescents’ criminal culpability. *Nature* **14**, 513–518. (doi:10.1038/nrn3509)
- Romer D. 2010 Adolescent risk taking, impulsivity, and brain development: implications for prevention. *Dev. Psychobiol.* **52**, 263–276. (doi:10.1002/dev.20442)
- Zuckerman M. 2007 *Sensation seeking and risky behavior*. Washington, DC: American Psychological Association.
- Kelley AE, Schochet T, Landry CF. 2004 Risk taking and novelty seeking in adolescence: introduction to part I. *Ann. NY Acad. Sci.* **1021**, 27–32. (doi:10.1196/annals.1308.003)
- Cross CP, Copping LT, Campbell A. 2011 Sex differences in impulsivity: a meta-analysis. *Psychol. Bull.* **137**, 97–130. (doi:10.1037/a0021591)
- Charness G, Gneezy U, Imas A. 2013 Experimental methods: eliciting risk preferences. *J. Econ. Behav. Organ.* **87**, 43–51. (doi:10.1016/j.jebo.2012.12.023)
- Schonberg T, Fox CR, Poldrack RA. 2011 Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends Cogn. Sci.* **15**, 11–19. (doi:10.1016/j.tics.2010.10.002)
- Moffitt TE *et al.* 2011 A gradient of childhood self-control predicts health, wealth, and public safety.

Data accessibility. Data available at <https://www.dirkwlff.org/#data>.

Authors’ contributions. D.U.W analysed data; R.H., D.U.W. and R.M. wrote the paper.

Competing interest. We have no competing interests.

Funding. This work was supported by the Swiss National Science Foundation [156172 to R.M.].

Acknowledgement. We are grateful to Deb Ain and Laura Wiles for editing the manuscript.

Endnotes

¹Admittedly, this binary distinction is a simplification. In our past work, we have further distinguished between self-reported propensity measures (assessing stated preferences) and self-reported frequency measures (tracking specific and observable behaviours [5]). Others have considered epidemiological data, such as crime or cause-specific mortality [17]; administrative data, such as arrests [24]; field data, such as property or health insurance data [1] and informant reports from relatives or acquaintances [25]. Most of the work on risk preference, however, rests on behavioural and self-report measures; our focus here is therefore on them.

²Samuelson [26] aimed to overcome what he criticized as the ‘discrediting of utility as a psychological concept’ within cardinal utility theory that cannot explain human behaviour because of its ‘circular sense, revealing its emptiness as even a construction’ (p. 61).

³The literature has adopted different ways of operationalizing the description–experience gap. One approach is to count how many of individuals’ actual choices in the description and experience condition, respectively, are consistent with the predicted choices, based on cumulative prospect theory’s (CPT) parameters (commonly using those derived by [52]). These parameters, derived from stated probabilities, embody overweighting of rare events; therefore, choices consistent with the predictions of CPT indicate a tendency to overweight rare events. When this definition is applied, a description–experience gap emerges when systematically fewer experience-based than description-based choices are correctly predicted.

- Proc. Natl Acad. Sci. USA* **108**, 2693–2698. (doi:10.1073/pnas.1010076108)
25. Roberts BW, Lejuez C, Krueger RF, Richards JM, Hill PL. 2014 What is conscientiousness and how can it be assessed? *Dev. Psychol.* **50**, 1315–1330. (doi:10.1037/a0031109)
 26. Samuelson PA. 1938 A note on the pure theory of consumer's behavior. *Economica* **5**, 61–71. (doi:10.2307/2548836)
 27. Samuelson PA. 1948 Consumption theory in terms of revealed preference. *Economica* **15**, 243–253. (doi:10.2307/2549561)
 28. Ellsberg D. 1961 Risk, ambiguity, and the Savage axioms. *Q. J. Econ.* **75**, 643–669. (doi:10.2307/1884324)
 29. Harrison GW, Rutström EE. 2008 Risk aversion in the laboratory. In *Risk aversion in experiments, research in experimental economics* (eds JC Cox, GW Harrison), vol. 12, pp. 41–196. Bingley, UK: Emerald.
 30. Holt CA, Laury SK. 2002 Risk aversion and incentive effects. *Am. Econ. Rev.* **92**, 1644–1655. (doi:10.1257/000282802762024700)
 31. Hacking I. 1975 *The emergence of probability*. Cambridge, UK: Cambridge University Press.
 32. Allais M. 1953 Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine [Rational man's behavior in the presence of risk: Critique of the postulates and axioms of the American school]. *Econometrica* **21**, 503–546. (doi:10.2307/1907921)
 33. Dohmen T, Huffman D, Schupp J, Falk A, Sunde U, Wagner GG. 2011 Individual risk attitudes: measurement, determinants, and behavioral consequences. *J. Eur. Econ. Assoc.* **9**, 522–550. (doi:10.1111/j.1542-4774.2011.01015.x)
 34. Blais A-R, Weber EU. 2006 A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgm. Decis. Making* **1**, 33–47.
 35. Marinelli N, Mazzoli C. 2011 An insight into suitability practice: is a standard questionnaire the answer? In *Bank strategy, governance and ratings* (ed. P Molyneux), pp. 217–245. London, UK: Palgrave Macmillan.
 36. Zuckerman M, Kolin EA, Price L, Zoob I. 1964 Development of a sensation-seeking scale. *J. Consult. Psychol.* **28**, 477–482. (doi:10.1037/h0040995)
 37. Dohmen T, Falk A, Golsteyn BH. H., Huffman D, Sunde U. 2017 Risk attitudes across the life course. *Econ. J.* **127**, F95–F116. (doi:10.1111/eoj.12322)
 38. Josef AK, Richter D, Samanez-Larkin GR, Wagner GG, Hertwig R, Mata R. 2016 Stability and change in risk-taking propensity across the adult lifespan. *J. Pers. Soc. Psychol.* **111**, 430–450. (doi:10.1037/pspp0000090)
 39. Benjamin DJ *et al.* 2012 The genetic architecture of economic and political preferences. *Proc. Natl Acad. Sci. USA* **109**, 8026–8031. (doi:10.1073/pnas.1120666109)
 40. Linnér RK *et al.* 2018 Genome-wide study identifies 611 loci associated with risk tolerance and risky behaviors. *bioRxiv* 261081. (doi:10.1101/261081)
 41. Weber EU, Blais A-R, Betz NE. 2002 A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *J. Behav. Decis. Making* **15**, 263–290. (doi:10.1002/bdm.414)
 42. Camerer C, Weber M. 1992 Recent developments in modeling preferences: uncertainty and ambiguity. *J. Risk Uncertain.* **5**, 325–370. (doi:10.1007/BF00122575)
 43. Gigerenzer G, Hertwig R, van den Broek E, Fasolo B, Katsikopoulos KV. 2005 'A 30% chance of rain tomorrow': how does the public understand probabilistic weather forecasts? *Risk Anal.* **25**, 623–629. (doi:10.1111/j.1539-6924.2005.00608.x)
 44. Hertwig R, Barron G, Weber EU, Erev I. 2004 Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* **15**, 534–539. (doi:10.1111/j.0956-7976.2004.00715.x)
 45. Wulff DU, Mergenthaler-Canseco M, Hertwig R. 2018 A meta-analytic review of two modes of learning and the description–experience gap. *Psychol. Bull.* **144**, 140–176. (doi:10.1037/bul0000115)
 46. Hertwig R, Hogarth RM, Lejarraga T. 2018 Experience and description: exploring two paths to knowledge. *Curr. Dir. Psychol. Sci.* **27**, 123–128. (doi:10.1177/0963721417740645)
 47. Barron G, Erev I. 2003 Small feedback-based decisions and their limited correspondence to description-based decisions. *J. Behav. Decis. Making* **16**, 215–233. (doi:10.1002/bdm.443)
 48. Hertwig R. 2015 Decisions from experience. In *The Wiley Blackwell handbook of judgment and decision making* (eds G Keren, G Wu), Vol. 1, pp. 239–267. Chichester, UK: Wiley Blackwell.
 49. Hertwig R, Erev I. 2009 The description–experience gap in risky choice. *Trends Cogn. Sci.* **13**, 517–523. (doi:10.1016/j.tics.2009.09.004)
 50. Rakow T, Newell BR. 2010 Degrees of uncertainty: an overview and framework for future research on experience-based choice. *J. Behav. Decis. Making* **23**, 1–14. (doi:10.1002/bdm.681)
 51. Cohen JD, McClure SM, Yu AJ. 2007 Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil. Trans. R. Soc. B* **362**, 933–942. (doi:10.1098/rstb.2007.2098)
 52. Tversky A, Kahneman D. 1992 Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertainty* **5**, 297–323. (doi:10.1007/BF00122574)
 53. Hertwig R. 2012 The psychology and rationality of decisions from experience. *Synthese* **187**, 269–292. (doi:10.1007/s11229-011-0024-4)
 54. Ludvig EA, Spetch ML. 2011 Of black swans and tossed coins: is the description–experience gap in risky choice limited to rare events? *PLoS ONE* **6**, e20262. (doi:10.1371/journal.pone.0020262)
 55. Regenwetter M, Robinson MM. 2017 The construct–behavior gap in behavioral decision research: a challenge beyond replicability. *Psychol. Rev.* **124**, 533–550. (doi:10.1037/rev0000067)
 56. Gonzalez C, Dutt V. 2011 Instance-based learning: integrating sampling and repeated decisions from experience. *Psychol. Rev.* **118**, 523–551. (doi:10.1037/a0024558)
 57. Camilleri AR, Newell BR. 2013 Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. *Prog. Brain Res.* **202**, 55–71. (doi:10.1016/B978-0-444-62604-2.00004-6)
 58. Hills TT, Hertwig R. 2010 Information search in decisions from experience: do our patterns of sampling foreshadow our decisions? *Psychol. Sci.* **21**, 1787–1792. (doi:10.1177/0956797610387443)
 59. Wulff DU, Hills TT, Hertwig R. 2015 Online product reviews and the description–experience gap. *J. Behav. Decis. Making* **28**, 214–223. (doi:10.1002/bdm.1841)
 60. Hawkins G, Camilleri A, Heathcote A, Newell BR, Brown SD. 2014 Modeling probability knowledge and choice in decisions from experience. *Proc. Annual Meeting of the Cognitive Science Society* 36. See <https://pub-jschol-prd.escholarship.org/uc/item/4mv5028f>.
 61. Kahneman D, Tversky A. 1984 Choices, values, and frames. *Am. Psychol.* **39**, 341–350. (doi:10.1142/9789814417358_0016)
 62. Birnbaum MH. 2008 New paradoxes of risky decision making. *Psychol. Rev.* **115**, 463–501. (doi:10.1037/0033-295X.115.2.463)
 63. Rehder B, Waldmann MR. 2017 Failures of explaining away and screening off in described versus experienced causal learning scenarios. *Mem. Cognit.* **45**, 245–260. (doi:10.3758/s13421-016-0662-3)
 64. Armstrong B, Spaniol J. 2017 Experienced probabilities increase understanding of diagnostic test results in younger and older adults. *Med. Decis. Making* **37**, 670–679. (doi:10.1177/0272989x17691954)
 65. Heilbronner SR, Hayden BY. 2016 The description–experience gap in risky choice in nonhuman primates. *Psychon. Bull. Rev.* **23**, 593–600. (doi:10.3758/s13423-015-0924-2)
 66. Caspi A *et al.* 2014 The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin. Psychol. Sci.* **2**, 119–137. (doi:10.1177/2167702613497473)
 67. Deary IJ. 2001 Human intelligence differences: a recent history. *Trends Cogn. Sci.* **5**, 127–130. (doi:10.1016/S1364-6613(00)01621-1)
 68. Highhouse S, Nye CD, Zhang DC, Rada TB. 2016 Structure of the Dospert: is there evidence for a general risk factor? *J. Behav. Decis. Making* **30**, 400–406. (doi:10.1002/bdm.1953)
 69. Mata R, Josef AK, Samanez-Larkin GR, Hertwig R. 2011 Age differences in risky choice: a meta-analysis. *Ann. NY Acad. Sci.* **1235**, 18–29. (doi:10.1111/j.1749-6632.2011.06200.x)
 70. Berg J, Dickhaut J, McCabe K. 2005 Risk preference instability across institutions: a dilemma. *Proc. Natl Acad. Sci. USA* **102**, 4209–4214. (doi:10.1073/pnas.0500333102)
 71. Stahl DO. 2018 Assessing the forecasting performance of models of choice. *J. Behav. Exp. Econ.* **73**, 86–92. (doi:10.1016/j.socec.2018.02.006)

72. Pedroni A, Frey R, Bruhin A, Dutilh G, Hertwig R, Rieskamp J. 2017 The risk elicitation puzzle. *Nat. Hum. Behav.* **1**, 803–809. (doi:10.1038/s41562-017-0219-x)
73. Vazire S, Carlson EN. 2010 Self-knowledge of personality: do people know themselves? *Soc. Pers. Psychol. Compass* **4**, 605–620. (doi:10.1111/j.1751-9004.2010.00280.x)
74. Schwarz N. 1999 Self-reports: how the questions shape the answers. *Am. Psychol.* **54**, 93–105. (doi:10.1037/0003-066X.54.2.93)
75. Haefffel GJ, Howard GS. 2010 Self-report: psychology's four-letter word. *Am. J. Psychol.* **123**, 181–188. (doi:10.5406/amerjpsyc.123.2.0181)
76. Duckworth AL, Kern ML. 2011 A meta-analysis of the convergent validity of self-control measures. *J. Res. Pers.* **45**, 259–268. (doi:10.1016/j.jrp.2011.02.004)
77. Galizzi MM, Navarro-Martinez D. 2018 On the external validity of social preference games: a systematic lab–field study. *Manage. Sci.* Advance online publication. (doi:10.1287/mnsc.2017.2908)
78. Eisenberg IW *et al.* 2017 Applying novel technologies and methods to inform the ontology of self-regulation. *Behav. Res. Ther.* **101**, 46–57. (doi:10.1016/j.brat.2017.09.014)
79. McNamara JM, Houston AI. 1992 Risk-sensitive foraging: a review of the theory. *Bull. Math. Biol.* **54**, 355–378. (doi:10.1007/BF02464838)
80. Mishra S. 2014 Decision-making under risk: integrating perspectives from biology, economics, and psychology. *Pers. Soc. Psychol. Rev.* **18**, 280–307. (doi:10.1177/1088868314530517)
81. Stephens D.W. 1981 The logic of risk-sensitive foraging preferences. *Anim. Behav.* **29**, 628–629. (doi:10.1016/S0003-3472(81)80128-5)
82. Bonsang E, Dohmen T. 2015 Risk attitude and cognitive aging. *J. Econ. Behav. Organ.* **112**, 112–126. (doi:10.1016/j.jebo.2015.01.004)
83. Kurath J, Mata R. 2018 Individual differences in risk taking and endogenous levels of testosterone, estradiol, and cortisol: a systematic literature search and three independent meta-analyses. *Neurosci. Biobehav. Rev.* **90**, 428–446. (doi:10.1016/j.neubiorev.2018.05.003)
84. Mata R, Josef AK, Hertwig R. 2016 Propensity for risk taking across the life span and around the globe. *Psychol. Sci.* **27**, 231–243. (doi:10.1177/0956797615617811)
85. van den Bos W, Hertwig R. 2017 Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Sci. Rep.* **7**, 40962. (doi:10.1038/srep40962)
86. Lichtenstein S, Slovic P. 1971 Reversals of preference between bids and choices in gambling decisions. *J. Exp. Psychol.* **89**, 46–55. (doi:10.1037/h0031207)
87. Chater N, Johansson P, Hall L. 2011 The non-existence of risk attitude. *Front. Psychol.* **2**, 303. (doi:10.3389/fpsyg.2011.00303)
88. Hermansson C. 2018 Can self-assessed financial risk measures explain and predict bank customers' objective financial risk? *J. Econ. Behav. Organ.* **148**, 226–240. (doi:10.1016/j.jebo.2018.02.018)
89. Coppola M. 2014 Eliciting risk-preferences in socio-economic surveys: how do different measures perform? *J. Socio-Econ.* **48**, 1–10. (doi:10.1016/j.socec.2013.08.010)
90. Friedman DR, Isaac RM, James D, Sunder S. 2014 *Risky curves: on the empirical failure of expected utility*. New York, NY: Routledge.