

Risk Preference: A View from Psychology

Rui Mata, Renato Frey, David Richter, Jürgen Schupp, and Ralph Hertwig

Investing in financial markets, engaging in criminal activity, or consuming recreational and possibly illicit drugs are examples of behaviors that involve trading-off potential costs and benefits associated with some degree of risk and uncertainty. Many psychologists aim to uncover the extent to which stable personality characteristics—psychological traits—account for why individuals differ in their appetite for risk and in their decision to engage in such behaviors. The endeavors of psychologists not only reflect an effort to understand human behavior per se, but also aim to better diagnose and prevent undesirable levels of risk taking, with the ultimate goal of improving the physical or mental health and the financial well-being of individuals and populations. In what follows, we use the term “risk preference” to refer to such a psychological trait (or collection of traits) and explore the extent to which both psychologists and economists can use it to explain individual differences in people’s appetite for risk.

Debates surrounding the nature of risk preference and its measurement have a long history in psychology and economics, and the number of discussion points

■ *Rui Mata is Professor for Cognitive and Decision Sciences and Renato Frey is a Researcher in Cognitive and Decision Sciences, both at the University of Basel, Basel, Switzerland. Mata and Frey are also Adjunct Researchers, Max Planck Institute for Human Development, Berlin, Germany. David Richter is Research Associate and Jürgen Schupp is the Director of the German Socio-Economic Panel Study, both at the German Institute for Economic Research, Berlin, Germany. Ralph Hertwig is Director of the Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany. Their email addresses are rui.mata@unibas.ch, renato.frey@unibas.ch, drichter@diw.de, jschupp@diw.de, and hertwig@mpib-berlin.mpg.de.*

† For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at <https://doi.org/10.1257/jep.32.2.155>

doi=10.1257/jep.32.2.155

is large (Bernoulli 1738 [1954]; Edwards 1954; Slovic 1964; Schonberg, Fox, and Poldrack 2011; Friedman, Isaac, James, and Sunder 2014). In psychology, risk preference is commonly defined as the propensity to engage in behaviors or activities that are rewarding yet involve some potential for loss, including substance use, or criminal activities that may be associated with considerable physical and mental harm to individuals (Steinberg 2013). In economics, risk preference more often refers to the tendency to engage in behaviors or activities that involve higher variance in returns, regardless of whether these represent gains or losses, and is often studied in the context of monetary payoffs involving lotteries (Harrison and Rutström 2008). Beyond such differences in definition and scope between fields, which we will not fully address, there are shared and unresolved conceptual and measurement issues overdue for consideration by both fields. We argue that psychology offers conceptual and analytic tools that can help advance the discussion on the nature of risk preference and its measurement in the behavioral sciences. We also provide an overview of strengths and weaknesses of two different measurement traditions of risk preferences that have coexisted in psychology and, to some extent, in economics: the revealed and stated preference traditions (Beshears, Choi, Laibson, and Madrian 2008; Appelt, Milch, Handgraaf, and Weber 2011; Charness, Gneezy, and Imas 2013). Lurking beneath these measurement aspects are broader conceptual issues. Let us briefly preview three before discussing them in more detail in the remainder of the article: temporal stability, convergent validity, and predictive validity.

Psychological traits, by definition, show some degree of *temporal stability*. Consequently, any theorizing about risk preference as a psychological trait must ask whether it shows a degree of stability over time that approximates what has been established for other major traits, such as intelligence, or, alternatively, is more similar to the stability of transitory psychological states, such as emotional states. Of course, no psychological trait is perfectly stable, and it may be subject to systematic variation as a function of specific contextual influences (Caspi, Roberts, and Shiner 2005). Such a view is compatible with our proposal that risk preference can both be seen as a stable psychological trait and yet show systematic and sizable changes as a function of specific life stages or momentary shocks (see also Schildberg-Hörisch, in this issue).

Convergent validity refers to the degree to which different measures of a psychological construct capture a common underlying characteristic or trait. Do measures of risk preference all capture a unitary psychological trait that is indicative of risky behavior across various domains, or do they capture various traits that independently contribute to risky behavior in specific areas of life, such as financial, health, and recreational domains (Weber, Blais, and Betz 2002; Highhouse, Nye, Zhang, and Rada 2017)? This need not be an either–or choice. For example, research on the trait of intelligence suggests that a single general factor can account for the largest share of variance (approximately 50 percent) in performance across many different tasks, with the rest of the variance being accounted for by more specific factors such as visual-spatial or logical-mathematical intelligence (Deary 2001). Similar results have been obtained for psychopathology: About 50 percent of variance in symptomatology is captured by a general factor, which is in line with the fact that about half of

individuals who meet diagnostic criteria for one disorder also meet diagnostic criteria for a second one (Caspi et al. 2014; Castellanos-Ryan et al. 2016). Critically, recent work on risk preference suggests that it may share the psychometric structure of such major psychological traits, by which over 50 percent of the systematic variance in measures of risk preference are accounted for by a general factor, with the remaining variance being shared among several additional specific factors (Frey, Pedroni, Mata, Rieskamp, and Hertwig 2017). Consequently, it may be important to consider the explanatory power of a general trait of risk preference in addition to more specific ones when accounting for individual differences in the appetite for risk.

Predictive validity refers to the extent to which a psychological trait has power in forecasting behavior. For example, intelligence and major personality traits, such as some of the Big Five traits (openness, conscientiousness, extraversion, agreeableness, neuroticism), have been shown to predict important life outcomes, such as academic and professional achievement (Schmidt and Hunter 2004; Richardson, Abraham, and Bond 2012). Such work suggests that it is important to examine the short- and long-term outcomes of risk preference—something that is still largely lacking in current psychological (and economic) research.

In what follows, we discuss the current empirical knowledge on risk preferences in light of these three arguments. However, first, we provide an overview of the revealed and stated preference measurement traditions, which have coexisted in both psychology and economics in the study of risk preferences. Without acknowledging their existence and understanding their somewhat difficult relation, it is hard to make concerted progress in research on risk preference.

Two Measurement Traditions

In his presidential address to the American Psychological Association, Lee Cronbach (1957), a towering figure of 20th century psychology, distinguished between two research streams that run through the history of the still young—and back then even younger—discipline of scientific psychology. One stream, he argued, is *experimental* psychology (see also Hertwig and Ortmann 2001). Its emphasis is on well-controlled experimental designs and on the goal of rigorously testing the influence of selected situational variables on behavior, cognition, and emotion, often using objective measures—such as overt choices and associated reaction times—as outcomes of interest. The other stream, *correlational* psychology, relies on observational and correlational designs to understand cross-situational and intra-individual consistency of the same behavior, cognition, and emotion, often with the aid of self-reports in response to standardized survey measures. Whereas experimenters' interest lies primarily in the impact of the variations they caused, the concern of correlators is with the (co)variation of individuals' behavior across naturally occurring situations.

Six decades later, the partition of psychological research into these two streams is still noticeable (Tracy, Robins, and Sherman 2009)—and perhaps nowhere more so than in research on the construct(s) of risk preference (Appelt et al. 2011; Frey et al.

2017). This distinction is also reflected in two major measurement approaches: one that mostly employs behavioral paradigms, and another that predominantly uses self-reports. These two broad approaches can also be identified, alongside others, in the economics literature (Beshears et al. 2008; Charness, Gneezy, and Imas 2013).¹

The behavioral stream in psychology focuses on understanding the cognitive or neural correlates of risk preference. This work often emphasizes the structural properties of tasks and environments that are associated with sometimes surprisingly different and even seemingly inconsistent behaviors (Kahneman and Tversky 1979; Mata, Josef, Samanez-Larkin, and Hertwig 2011). For example, a long tradition in both economics and psychology uses choices between lotteries to understand how individuals deal with gains and losses or specific types of incentive structures (Kahneman and Tversky 1979; Holt and Laury 2002). This type of research is alive and well in its somewhat splendid isolation—an issue to which we return shortly. For example, recent experimental efforts have tried to understand the *description–experience gap* that arises from differences in the presentation format of risk information (Hertwig and Erev 2009). For example, the numerical description of risks (“stated probabilities”) in canonical lottery tasks gives rise to choices indicative of overweighting of small probabilities, but sequential experience of risk first-hand through sampling of outcomes is associated with choices as if people underweight small probabilities (Wulff, Mergenthaler-Canseco, and Hertwig 2018). A large swath of research now aims to identify the neural basis of choice in such experience-based and description-based paradigms using functional neuroimaging and other neuroscientific methods (Glimcher and Fehr 2014; Knutson and Huettel 2015). Researchers from this approach often focus on uncovering the psychological processes underlying choices in a specific behavioral paradigm, but often with little or no investigation of how such processes generalize across paradigms and time.

Studies using self-report measures seek to elicit *stated preferences* in response to hypothetical or real-world behaviors. For example, respondents may be asked to rate themselves on a rating scale with opposite poles being “not at all willing to take risks” and “very willing to take risks,” or express the likelihood of engaging in some risky behavior—“How likely would you be to go white-water rafting at high water in the spring?” A growing body of work on risk preference builds primarily on findings from either single-item (Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner 2011) or multiple-item self-report measures of risk preference (Blais and Weber 2006). For example, this type of data has been used to study stable individual characteristics, such

¹For the sake of completeness, it should be noted that there are other approaches to studying and measuring risk preference in both psychology and economics. Frey et al. (2017) distinguished between behavioral measures (assessing revealed preferences), self-reported propensity measures (assessing stated preferences), and self-reported frequency measures (tracking specific and observable behaviors). Other approaches include the use of epidemiological data from population statistics, such as crime or cause-specific mortality (Steinberg 2013), actual behavior as captured from administrative or survey data (Moffitt et al. 2011), or observer reports from relatives or acquaintances (Roberts, Lejuez, Krueger, Richards, and Hill 2014). However, the bulk of work on risk preference rests on behavioral and self-report measures, so we focus on those here.

as the genetic basis of risk preference (see also Benjamin et al. 2012; Beauchamp, Cesarini, and Johannesson 2017) as well as to uncover cohort (Malmendier and Nagel 2011; Dohmen, Falk, Golsteyn, Huffman, and Sunde 2017), life span (Josef, Richter, Samanez-Larkin, Wagner, Hertwig, and Mata 2016; Dohmen et al. 2017), and momentary (Browne, Jaeger, Richter, and Steinorth 2016) changes in risk preference. Importantly, such self-report preference measures are now included in a number of panel assessments, such as the German Socio-Economic Panel (Wagner, Frick, and Schupp 2007), the US Health and Retirement Survey (Fisher, Gideon, Hsu, and McFall 2011), the British Household Panel (Galizzi, Machado, and Miniaci 2016), the Swedish Screening Across the Lifespan Twin survey (Beauchamp, Cesarini, and Johannesson 2017), the Swiss Household Panel (Mamerow, Frey, and Mata 2016), and the Household, Income and Labour Dynamics in Australia (HILDA) Survey (Clark and Lisowski 2017). These panel studies are important data troves for revealing more about the associates and determinants of risk preference.

Both behavioral and self-report measures of risk preference have been subject to criticism. For example, some have voiced concern about the lack of generalizability across behavioral elicitation methods (Friedman et al. 2014). There is also a fair amount of skepticism in both economics (Beshears et al. 2008) and psychology (Haeffel and Howard 2010) about self-reports representing little more than “cheap talk.” In our view, the relative strengths and weaknesses of the two measurement approaches as well as their possible links should be determined empirically. Unfortunately, and echoing Cronbach’s (1957) diagnosis of psychologists’ firm commitment to either one or the other methodology and associated theoretical constructs, the behavioral and the self-report approaches to measuring risk preference usually exist side-by-side with little or no empirical or theoretical integration.

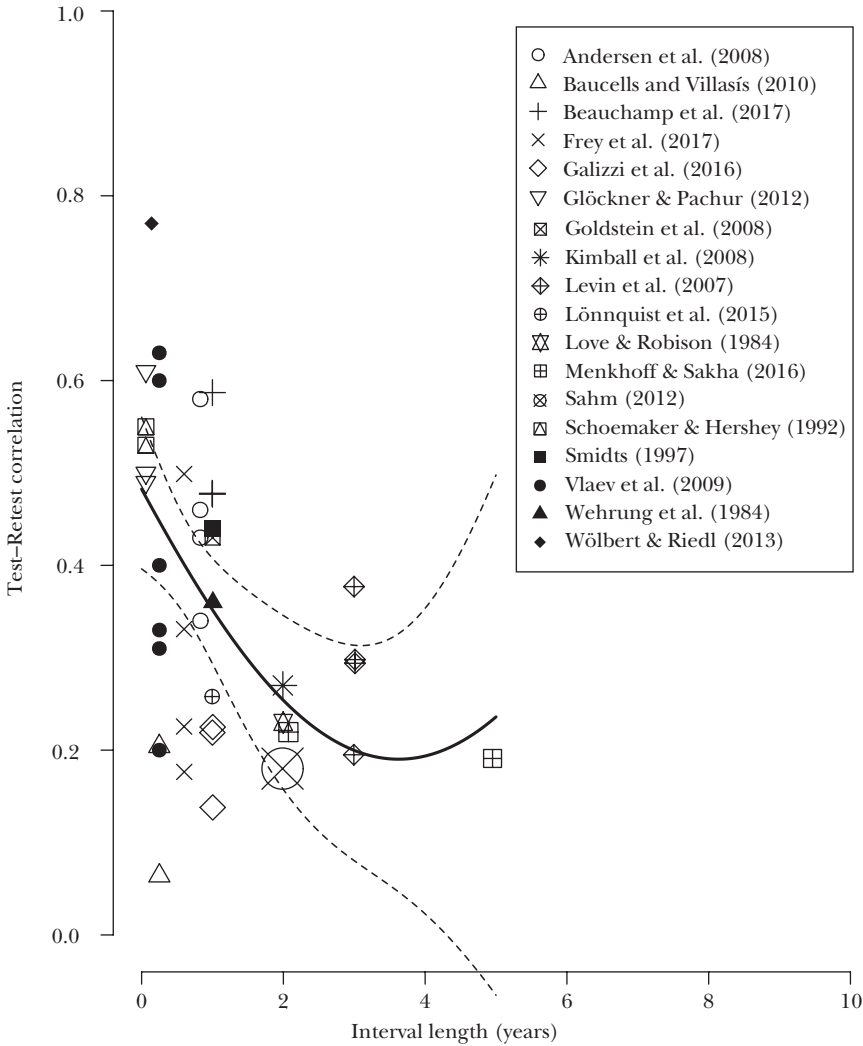
Next, we turn to some of the work produced by the two approaches to studying risk preference, emphasize the strengths and weaknesses of both, and address implications for a more general theory of risk preference. In particular, we provide some evidence that self-report measures represent stable indicators of risk preference whereas widely used behavioral measures do not—and possibly as a consequence, there is often little agreement between the two. This realization has at least one important implication for psychologists and economists studying risk preference: It suggests that measures of risk preference cannot be used interchangeably when predicting outcomes of interest.

Temporal Stability

Do revealed (behavioral) and stated (self-report) risk preferences show similar levels of temporal stability? For an admittedly preliminary answer, we took advantage of narrative reviews of past work (Chuang and Schechter 2015) and drew on our knowledge of the literature to identify published findings and datasets that allowed us to compute test–retest reliability of revealed and stated risk preferences. Specifically, we identified studies reporting test–retest reliability of choices between

Figure 1
Meta-analysis of Test–Retest Stability of Risk Preferences

A: Choices between Lotteries

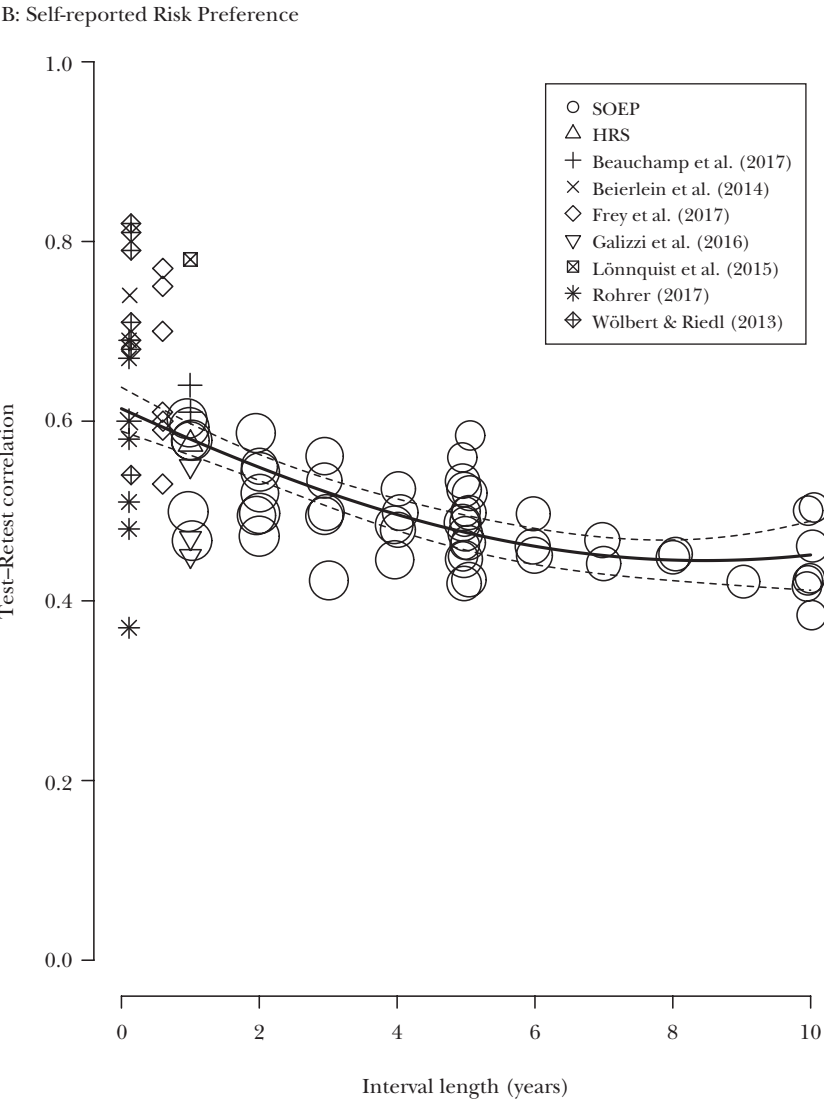


(Continued on next page)

monetary lotteries (for example, Harrison and Rutström 2008), as well as studies and datasets reporting test–retest reliability of self-report items, with those items probing the propensity to take risks either in general or in specific domains of life, such as financial, health, and social domains (for example, Dohmen et al. 2011).

Figure 1 depicts the meta-analytic scatterplots of test–retest correlations for choices between lotteries (Figure 1A) and for self-reported risk preference (Figure 1B). The test–retest correlations help assess the extent to which the same

Figure 1 (Continued)
Meta-analysis of Test–Retest Stability of Risk Preferences



Note: Figure 1 presents the meta-analytic scatterplots of test–retest correlations for choices between lotteries (Panel A) and self-reported risk preference (Panel B). Symbols represent correlations between two measurement occasions obtained from published literature (references provided in the figure legends) and our own calculations (German Socio-Economic Panel, SOEP; American Health and Retirement Survey, HRS). Note that a small amount of jitter was added to each point to better distinguish points at the same interval length. The size of each point is proportional to the inverse variance (larger symbols = more precision). The solid line represents the weighted regression line including a linear and a quadratic term for interval length from a random effects meta-analysis (dashed lines correspond to 95 percent confidence intervals). We conducted the analyses using the package metafor for R (Viechtbauer 2010). Data and code are provided online with the article at the journal website, at <https://www.aeaweb.org/journals/jep>.

rank-ordering of individuals is preserved across two waves. To our knowledge, no data are available about the temporal stability for choices between lotteries with retest intervals longer than five years. Data on temporal stability of up to 10 years are available for self-report measures, albeit stemming mostly from one source, the German Socio-Economic Panel (Wagner, Frick, and Schupp 2007). Our analysis suggests that after five years, the measures taken from choices between lotteries show test–retest correlations of around .2 (although there is considerable uncertainty around that estimate). In contrast, the corresponding correlations for self-report are around .5 and these values do not seem to decline much across a 10-year period. Indeed, the level of stability found for self-report measures of risk preference is only slightly below the 10-year stability estimates for major personality traits, such as the Big Five, which are estimated at about .6, and shows greater stability than measures of life satisfaction, self-esteem, and affect, which are estimated to range between .35 and .4 for a 10-year period (Anusic and Schimmack 2016).

A potential criticism of the meta-analysis for measures of revealed preference is that it relies on choices between lotteries, and such choices may be perceived as artificial, therefore failing to engage participants. However, we have examined test–retest reliability of other prominent behavioral risk preference measures, including measures designed to be more engaging, such as the Balloon Analogue Risk Task (Lejuez et al. 2002) or the Columbia Card Task (Figner, Mackinlay, Wilkening, and Weber 2009). Such measures show low levels of test–retest reliability similar to those found using choices between lotteries across a delay of six months (Frey et al. 2017).

Does high temporal stability of risk preferences for individuals, at least for stated preferences, mean that there are no systematic changes within individuals over shorter or longer time scales? No. Research on personality suggests that high temporal stability in differences between individuals across long intervals is compatible with population mean-level changes in psychological traits (Roberts and DelVecchio 2000). Stability and change are compatible because mean-level changes—say, changes across the lifespan—represent average patterns affecting many or all individuals in the population, whereas test–retest reliability captures preservation of the relative *rank-ordering* of individuals, regardless of mean-level differences. This point may be easier to appreciate with an example. Intelligence is one of the most stable constructs known to psychology because of evidence of preserved rank-order stability (within a cohort) across decades (Deary 2001). However, intelligence can show dramatic and systematic changes as a function of momentary shocks, such as sleep deprivation (Lim and Dinges 2010) as well as long-term changes across the life span, including considerable decline in fluid components, such as reasoning and memory, with aging (Baltes, Staudinger, and Lindenberger 1999; Lindenberger 2014). Consistent with the concurrent presence of stability and change, we and others have found high test–retest reliability (Josef et al. 2016) *as well as* systematic mean-level reductions in risk-taking propensity with age in longitudinal examinations of self-reported risk-taking propensity (Josef et al. 2016; Dohmen et al. 2017).

One outstanding issue concerning individual and age-related changes in stated preferences is the extent to which they are indicative of “real” changes as opposed to mere changes in individuals’ use of reference points across time. A similar issue has been raised in the domain of subjective well-being, where some have argued that age differences in self-reported measures may represent different benchmarks or reference classes (Weimann, Knabe, and Schöb 2015). Presently, we cannot offer a satisfactory response to this possible objection. Ideally, one would tackle this issue by using measures that are robust to this criticism, such as self-report measures that provide a relatively stable referential context (for example, “how risk taking are you relative to those of your age?”) or, of course, behavioral measures in which reference points can be firmly and transparently established and systematically varied.

To summarize, risk preference measured from stated preferences emerges as a construct with considerable temporal stability, although revealed preference measures do not show such stability. Moderate rank-order stability in stated risk preferences is accompanied by sizable mean-level differences across the life span as well as significant variation within individuals. Consequently, the evidence suggests that present and future theories of risk preference need to account for both stable differences between individuals as well as systematic variation within individuals.

Convergent Validity

A key question in psychological research on risk preference has been whether it can be thought of as a domain-general tendency (similar to a general factor of intelligence, *g*, affecting behaviors implicating intelligence across many diverse contexts), or whether it should be construed as a multidimensional or domain-specific construct, with specific tendencies regarding wealth, health, or social exchange, to name just a few (for example, Slovic 1964; Weber, Blais, and Betz 2002). One way to approach this question empirically is to ask whether different measures of risk preference such as behavioral and self-report measures speak with one voice and converge in what they suggest about the individual.

Several studies on issues unrelated to risk have found that differences in experimental design can make a very large difference in behavioral patterns: for example, Berg, Dickhaut, and McCabe (2005) found large variations in behavioral patterns in laboratory experiments using different economic institutions, and Hertwig and Erev (2009) have found systematic differences and even preference reversals depending on whether risk information was described or experienced through repeated sampling. Further, the reported correlations between measures of risk preference are typically low (Dohmen et al. 2011; Galizzi, Machado, and Miniaci 2016). Such results cast doubt on the convergent validity of established risk preference measures. In what follows, we detail our recent efforts to assess the convergent validity of risk preference measures. We find a serious gap between different methods of eliciting risk preferences; in particular, we find a divide between stated (self-report) and

revealed (behavioral) preference measures, as well as among different behavioral measures.

First, in a study with 1,507 participants who completed a comprehensive battery of 39 risk preference measures—including a range of stated and revealed preference measures—we found that correlations between measures from the revealed and stated preference traditions were weak ($r = 0.06$; Frey et al. 2017). Moreover, the correlations among the nine different behavioral measures were substantially weaker ($r = 0.08$) than those among the 29 self-report measures ($r = 0.20$), even though the latter intentionally capture risk preference in diverse domains, such as financial, health, recreational, and social. The correlations between behavioral measures were not increased when parameters from specific decision models, such as expected utility theory or cumulative prospect theory, were used to describe individual's choices (Pedroni, Frey, Bruhin, Dutilh, Hertwig, and Rieskamp 2017). We also conducted a psychometric analysis using a bifactor model that directly accounts for shared variance across all measures with a single factor, leaving any residual variance to be captured by yet other specific, orthogonal factors. The bifactor analysis suggested that a general risk preference factor accounts for over 60 percent of the explained variance across measures, with the remaining variance captured by more domain-specific factors. Crucially, though this general factor explained substantial variance across self-report measures, it did not generalize to the behavioral measures. Overall, our psychometric analysis suggests that there is a large shared component that can be thought of as a general factor of risk preference bridging different domains of life that is captured from self-report (albeit not behavioral) measures. The idea of a general risk preference is in line with the robust observation that major psychological traits account for large portions of variance in subjective reports or behavior (Deary 2001; Caspi et al. 2014).

Second, we recently conducted a study on the gap between risk preference measures and its implications for understanding individual, sex, and age differences in risk preference, using the Innovation Sample of the German Socio-Economic Panel (Richter and Schupp 2015). In this study, we used different elicitation methods to survey a relatively large, age-heterogeneous, representative sample of the population, which ensures considerable variance in the outcomes of interest. Specifically, 951 individuals between 18 and 80 years of age were asked to complete different measures, including self-report measures of risk-taking propensity as well as incentivized behavioral measures of risk taking, involving decisions based on either described or experienced risk (Frey, Richter, Schupp, Hertwig, and Mata 2018). We were thus able to analyze the convergent validity of the three different measure types. Our findings are similar to past work on the description–experience gap, which suggests a gap in choice behavior between the measures involving the same lottery choices but presented in description mode or in experienced mode (Hertwig and Erev 2009; Wulff et al. 2018). Furthermore, we observed a gap between behavioral and self-report measures in their intercorrelations and their covariates. More precisely, the self-report, but not the behavioral measures, show the common patterns of sex and age differences identified in previous work, whereby males show higher

levels of risk-taking propensity relative to females, and younger adults show higher levels of risk-taking propensity relative to older adults (Josef et al. 2016; Mata, Josef, and Hertwig 2016). These data suggest not only a separation between self-reported and revealed preference measurements, but also systematic differences in how they relate to some demographic covariates.

Third, we have conducted several other studies that show that different behavioral measures also do not coalesce in providing a similar picture of age differences, which is potentially a result of the differential cognitive demands they impose (Mata et al. 2011; Frey, Mata, and Hertwig 2015; Mamerow et al. 2016). In a meta-analysis, we found that those behavioral measures of risk preference that involve considerable learning and memory demands are more likely to indicate large age differences in risk preferences (Mata et al. 2011). Specifically, whether older adults tend to be more risk-seeking relative to younger adults, or vice versa, is likely to depend on the architecture of the choice task. For instance, older adults appear as if they seek more risk, relative to younger adults, whenever learning is necessary to overcome a task-specific anchor to choose a seemingly attractive but ultimately disadvantageous risky option. These results suggest one cause for the gap *between* revealed and stated preferences and even *within* revealed preferences. Revealed preferences are derived from measures that enlist processes that are also subject to cognitive or learning abilities and thus to inter- and intra-individual (during a life span) variations on those processes (for additional discussion of the role of cognitive abilities see the article by Dohmen, Falk, Huffman, and Sunde in this symposium).

To summarize, at present, there appears to be little hope for establishing a clear link between self-report and behavioral measures of risk preference, not only because measures from the two traditions do not correlate with each other, but also because revealed preference measures, that is, behavioral measures, fail to converge. Nevertheless, extant work suggests that stated preferences partly derive from a general risk preference component that accounts for a large portion of variance across life domains. As discussed in the next section, whether stated or revealed preference measures provide a better account of individuals' propensity for risk should be judged in light of prospective studies involving predictive validity of real-world behavior.

Predictive Validity

Real-world financial institutions such as banks and insurance companies have shown little use for revealed risk preference measures when recommending their products—perhaps because of the surprisingly limited predictive validity of utility and risk constructs obtained from revealed preference measures for real-world choices (Friedman et al. 2014). Unfortunately, there are few studies in the literature involving the measurement of risk preference to predict *objective* measures of real-world outcomes. Those few studies suggest, first, that self-reports and informant reports, assessing risk preference or related constructs, do have considerable

predictive validity for real-world outcomes such as teenage pregnancy, drug use, or financial security, even when controlling for other factors such as intelligence or socioeconomic status (Moffitt et al. 2011; Caspi et al. 2016; Beauchamp, Cesarini, and Johannesson 2017). Second, self- and informant reports are potentially more powerful than behavioral measures in this regard (White, Moffitt, Avshalom, Jørglum, Needles, and Stouthamer-Loeber 1994). In addition, interventions that have targeted specific at-risk groups identified through self-report measures of related constructs show promising results (Conrod et al. 2013), whereas the complementary evidence for the power of behavioral measures is still lacking.

To summarize, the current scant evidence suggests no advantage of revealed (behavioral) over stated preference measures in predicting real-world outcomes. While there are some promising results concerning the predictive validity of stated risk preference, data concerning the predictive validity of behavioral measures and comparisons between self-report and behavioral approaches are sorely needed. Clearly, more prospective longitudinal designs are required for both measurement paradigms. Such studies are difficult, time-consuming, and expensive to conduct. Unfortunately, long-standing panels, such as the German Socio-Economic Panel (SOEP), are not, at this time, equipped with psychometrically sound behavioral measures (for example, measures with satisfactory test-retest stability, batteries exhibiting convergent validity), nor objectively measured criterion variables (for example, credit reports, drug tests from biological samples) to permit fast progress in this regard. However, there is some work that links risk preference data from existing surveys to administrative data, such as education or income (for example, Beauchamp et al. 2017), and we hope and expect that more will follow.

A Look Ahead

Risk preference, when measured through stated, self-reported preferences, displays trait-like characteristics, such as high temporal stability across years and high convergent validity between different measurement instruments spanning different life domains. Furthermore, stated preferences seem to show significant predictive validity for a number of economic and health outcomes, dispelling the notion of self-assessments as simply “cheap talk.” However, the picture emerging from studies using revealed (behavioral) preference measures is less promising, with problems of poorer temporal stability, confounds related to high demands on learning, memory, or numeracy skills, and a relative lack of evidence concerning their predictive validity.

Many important phenomena in research on risk preference are still insufficiently understood. What explains the lack of convergence between stated and revealed preference measures? Why do revealed preference measures display so little convergent validity among themselves? What is the relative predictive validity of stated relative to revealed preference measures? In light of the fundamental nature of such questions, we hope that psychologists and economists team up to

conduct the necessary research to address them. We should emphasize that the debate on bridging the divide between different measures, such as the self-report and behavioral measures of risk preference, is not unlike that taking place in economics concerning the link between subjective and objective measures of well-being (Deaton and Stone 2013). To harvest the potential of these constructs and their value for actual policy-making, they need to still be better operationalized, measured, and understood.

Looking ahead, we identify two main avenues for future work on the study of risk preference. First, we hope to have helped to convince researchers interested in risk preference to undertake the painstaking task of examining the temporal stability, convergent validity, and predictive validity of their favorite measures. A time-honored tradition, such as relying on choices between monetary gambles, cannot substitute for this foundational work. Our own goal for future work is to develop and study a toolbox of measures to assess their strengths: perhaps some measures may be better at gauging a trait-like and domain-general component of risk preference, whereas others may be better suited to gauge domain-specific components. For example, it is possible that some behavioral measures may be better in simulating the specific incentive structure and choice architecture of a real-world context for which behavior is to be predicted. One interesting avenue toward a toolbox and taxonomy of risk preference measures is theory-driven task construction and decomposition using computational or neural methods that can disentangle risk preference from cognitive demands or other individual characteristics (Wallsten, Pleskac, and Lejuez 2005; Helfinstein et al. 2014). However, we suspect that computational and neural methods offer no panacea for the lack of temporal stability and convergent validity of the currently available behavioral measures (Frey et al. 2017; Pedroni et al. 2017).

Second, we need to make conceptual progress by addressing the psychological primitives or traits underlying individual differences in the appetite for risk. There is some agreement in the psychological literature about the existence of a few major psychological traits, such as a general factor of intelligence, *g*, and a few basic dimensions of personality (as one example, extraversion). However, there are still ongoing debates about distinctions within such constructs. In intelligence research, some lines of research focus on a general factor (Deary 2001) whereas others investigate specific facets such as the distinction between crystallized versus fluid intelligence (Baltes et al. 2007). Similarly, in the field of personality there are ongoing debates about whether to distinguish one, two, five, or yet more dimensions of personality (for example, Block 2010). The place for risk preference in this uncertain “periodic table” of psychological elements is yet unclear. Psychology has a tradition of introducing new constructs without full concern for their conceptual or empirical distinction. In this context, risk preference, sensation-seeking, impulsivity, self-control, grit, will-power, self-regulation, or conscientiousness are only some of the monikers that psychologists have introduced to explain individual differences in the appetite for risky activities, for example drug use, crime, and financial investment (Cross, Copping, and Campbell 2011; Roberts, Lejuez, Krueger, Richards, and

Hill 2014; Sharma, Markon, and Clark 2014). In line with the notion that psychological traits are general, we suspect that such labels characterize largely the same trait, and our empirical work suggests considerable overlap between such constructs (Frey et al. 2017). In practice, empirical studies investigating the temporal stability, and convergent and predictive validity of such different constructs and their respective measures will be fundamental in making conceptual progress. Psychology is already moving in that direction by initiating studies directly aimed at uncovering the amount of variance shared by measures originally proposed in the context of different traits (Eisenberg et al. 2018; Frey et al. 2017). The results of this work promise to be of immediate practical relevance to all behavioral scientists interested in determining how many and what kind of risk preference measures to include in their studies and models.

To conclude, risk preference, at least when measured through stated (self-reported) preferences, may be thought of as a moderately stable, general psychological trait, and, thus, an important variable to consider in psychological and economic theories and policy-making. Nevertheless, the measurement of risk preference needs more attention, and the usefulness of behavioral measures to uncover a stable psychological risk preference trait seems, at this time, surprisingly limited. Future research on risk preference needs to develop and deploy both stated (self-report) and revealed (behavioral) risk preference measures in prospective longitudinal studies in order to uncover their convergent and differential predictive validity for important economic and other life outcomes. In the meantime, behavioral scientists should be aware that not all measures of risk preference are created equal.

■ *We are grateful to the editors Gordon Henson, Enrico Moretti, and Timothy Taylor for their feedback; to Ulrike Malmendier for helpful discussions on the topic; to Anne Albrecht for help with coding studies for the meta-analyses; and to Laura Wiles and Ann Norman for editing the manuscript. This work was partly funded by a grant to the first author from the Swiss National Science Foundation (156172; <http://p3.snf.ch/project-156172>).*

References

- | | |
|---|---|
| <p>Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström. 2008. "Lost in State Space: Are Preferences Stable?" <i>International Economic Review</i>. 49(3): 1091-112.</p> | <p>Anusic, Ivana, and Ulrich Schimmack. 2016. "Stability and Change of Personality Traits, Self-Esteem, and Well-Being: Introducing the Meta-Analytic Stability and Change Model of</p> |
|---|---|

Retest Correlations." *Journal of Personality and Social Psychology* 110(5): 766–81.

Appelt, Kirstin C., Kerry F. Milch, Michel J. J. Handgraaf, and Elke U. Weber. 2011. "The Decision Making Individual Differences Inventory and Guidelines for the Study of Individual Differences in Judgment and Decision-Making Research." *Judgment and Decision Making* 6(3): 252–62.

Baltes, Paul B., Ursula M. Staudinger, and Ulman Lindenberger. 1999. "Lifespan Psychology: Theory and Application to Intellectual Functioning." *Annual Review of Psychology* 50: 471–507.

Baucells, Manel, and Antonio Villasís. 2010. "Stability of Risk Preferences and the Reflection Effect of Prospect Theory." *Theory and Decision* 68(1–2): 193–211.

Beauchamp, Jonathan P., David Cesarini, and Magnus Johannesson. 2017. "The Psychometric and Empirical Properties of Measures of Risk Preferences." *Journal of Risk and Uncertainty* 54(3): 203–37.

Beierlein, Constanze, Anastassiya Kovaleva, Christoph J. Kemper, and Beatrice Rammstedt. 2014. "Eine Single-Item-Skala zur Erfassung von Risikobereitschaft: Die Kurzskaala Risikobereitschaft-1 (R-1)" [A Single Item Scale for the Measurement of Proneness to Risk: The Short Scale Risk Proneness-1 (R-1)]. (GESIS Working Papers 2014|34) Cologne, Germany: GESIS.

Benjamin, Daniel J., David Cesarini, Matthijs J. H. M. van der Loos, Christopher T. Dawes, Philipp D. Koellinger, Patrik K. E. Magnusson, Christopher F. Chabris, Dalton Conley, David Laibson, Magnus Johannesson, and Peter M. Visscher. 2012. "The Genetic Architecture of Economic and Political Preferences." *PNAS* 109(21): 8026–31.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 2005. "Risk Preference Instability across Institutions: A Dilemma." *PNAS* 102(11): 4209–14.

Bernoulli, Daniel. 1738 [1954]. "Exposition of a New Theory on the Measurement of Risk." *Econometrica* 22(1): 23–36.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian. 2008. "How Are Preferences Revealed?" *Journal of Public Economics* 92(8–9): 1787–94.

Blais, Ann-Renée, and Elke U. Weber. 2006. "A Domain-Specific Risk-Taking (DOSPERT) Scale for Adult Populations." *Judgment and Decision Making* 1(1): 33–47.

Block, Jack. 2010. "The Five-Factor Framing of Personality and Beyond: Some Ruminations." *Psychological Inquiry* 21(1): 2–25.

Browne, Mark J., Verena Jaeger, Andreas Richter, and Petra Steinorth. 2016. "Family Transitions and Risk Attitude." MRIC Working Paper 32, Munich Risk and Insurance Center.

Caspi, Avshalom, Renate M. Houts, Daniel W. Belsky, Sidra J. Goldman-Mellor, HonaLee Harrington, Salomon Isreal, Madeline H. Meier, Sandhya Ramrakha, Idan Shalev, Richie Poulton, Terrie E. Moffitt. 2014. "The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders?" *Clinical Psychological Science* 2(2): 119–37.

Caspi, Avshalom, Renate M. Houts, Daniel W. Belsky, Hona Lee Harrington, Sean Hogan, Sandhya Ramrakha, Richie Poulton, and Terrie E. Moffitt. 2016. "Childhood Forecasting of a Small Segment of the Population with Large Economic Burden." *Nature: Human Behaviour* 1, Article 0005.

Caspi, Avshalom, Brent W. Roberts, and Rebecca L. Shiner. 2005. "Personality Development: Stability and Change." *Annual Review of Psychology* 56: 453–84.

Castellanos-Ryan, Natalie, et al. 2016. "The Structure of Psychopathology in Adolescence and Its Common Personality and Cognitive Correlates." *Journal of Abnormal Psychology* 125(8): 1039–52.

Charness, Gary, Uri Gneezy, and Alex Imas. 2013. "Experimental Methods: Eliciting Risk Preferences." *Journal of Economic Behavior and Organization* 87: 43–51.

Chuang, Yating, and Laura Schechter. 2015. "Stability of Experimental and Survey Measures of Risk, Time, and Social Preferences: A Review and Some New Results." *Journal of Development Economics* 117: 151–70.

Clark, William A. V., and William Lisowski. 2017. "Prospect Theory and the Decision to Move or Stay." *PNAS* 114(36): E7432–E77440.

Conrod, Patricia J., Maeve O'Leary-Barrett, Nicola C. Newton, Lauren Topper, Natalie Castellanos-Ryan, Clare Mackie, and Alain Girard. 2013. "Effectiveness of a Selective, Personality-Targeted Prevention Program for Adolescent Alcohol Use and Misuse: A Cluster Randomized Controlled Trial." *JAMA Psychiatry* 70(3): 334–9.

Cronbach, Lee J. 1957. "The Two Disciplines of Scientific Psychology." *American Psychologist* 12(11): 671–84.

Cross, Catharine P., Lee T. Copping, and Anne Campbell. 2011. "Sex Differences in Impulsivity: A Meta-Analysis." *Psychological Bulletin* 137(1): 97–101.

Deary, Ian J. 2001. "Human Intelligence Differences: A Recent History." *Trends in Cognitive Sciences* 5(3): 127–30.

Deaton, Angus, and Arthur A. Stone. 2013. "Two Happiness Puzzles." *American Economic Review* 103(3): 591–7.

Dohmen, Thomas, Armin Falk, Bart H. H. Golsteyn, David Huffman, and Uwe Sunde. 2017. "Risk Attitudes Across the Life Course." *Economic*

Journal 127(605): F95–F116.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. “Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences.” *Journal of the European Economic Association* 9(3): 522–50.

Edwards, Ward. 1954. “The Theory of Decision Making.” *Psychological Bulletin* 51(4): 380–417.

Eisenberg, Ian W., et al. 2018. “Applying Novel Technologies and Methods to Inform the Ontology of Self-Regulation.” *Behaviour Research and Therapy*, 101: 46–57.

Figner, Bernd, Rachael J. Mackinlay, Friedrich Wilkening, and Elke U. Weber. 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(3): 709–30.

Fisher, Gwenith, Mike Gideon, Joanne W. Hsu, and Brooke Helpie McFall. 2011. “Cognitive Economics Study: Development and Methodology.” November. Cognitive Economics Project, University of Michigan, Ann Arbor, USA. <http://ebp-projects.isr.umich.edu/CogEcon/docs/dev+meth.pdf>.

Frey, Renato, Rui Mata, and Ralph Hertwig. 2015. “The Role of Cognitive Abilities in Decisions from Experience: Age Differences Emerge as a Function of Choice Set Size.” *Cognition* 142: 60–80.

Frey, Renato, Andreas Pedroni, Rui Mata, Jörg Rieskamp, and Ralph Hertwig. 2017. “Risk Preference Shares the Psychometric Structure of Major Psychological Traits.” *Science Advances* 3(10): e1701381.

Frey, Renato, David Richter, Jürgen Schupp, Ralph Hertwig, and Rui Mata. 2018. “Identifying Correlates of Risk Preference: A Systematic Approach Using Specification Curve Analysis” Unpublished paper.

Friedman, Daniel, R. Mark Isaac, Duncan James, and Shyam Sunder. 2014. *Risky Curves: On the Empirical Failure of Expected Utility*. New York: Routledge.

Galizzi, Matteo M., Sara R. Machado, and Raffaele Miniaci. 2016. “Temporal Stability, Cross-Validity, and External Validity of Risk Preferences Measures: Experimental Evidence from a UK Representative Sample.” Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2822613>.

Glimcher, Paul W., and Ernst Fehr. 2014. *Neuroeconomics: Decision Making and the Brain*, 2nd edition. San Diego, USA: Academic Press, Elsevier.

Glöckner, Andreas, and Thorsten Pachur. 2012. “Cognitive Models of Risky Choice: Parameter Stability and Predictive Accuracy of Prospect Theory.” *Cognition* 123(1): 21–32.

Goldstein, Daniel G., Eric J. Johnson, and William F. Sharpe. 2008. “Choosing Outcomes versus Choosing Products: Consumer-Focused Retirement Investment Advice.” *Journal of Consumer Research* 35(3): 440–56.

Haefffel, Gerald J., and George S. Howard. 2010. “Self-Report: Psychology’s Four-Letter Word.” *American Journal of Psychology* 123(2): 181–88.

Harrison, Glenn W., and E. Elisabet Rutström. 2008. “Risk Aversion in the Laboratory.” *Research in Experimental Economics*, Vol. 12: *Risk Aversion in Experiments*, edited by James C. Cox and Glenn W. Harrison, pp. 41–196. Bingley, UK: Emerald.

Helfinstein, Sara M., Tom Schonberg, Eliza Congdon, Katherine H. Karlsgodt, Jeanette A. Mumford, Fred W. Sabb, Tyrone D. Cannon, Edythe D. London, Robert M. Bilder, and Russell A. Poldrack. 2014. “Predicting Risky Choices from Brain Activity Patterns.” *PNAS* 111(7): 2470–75.

Hertwig, Ralph, and Ido Erev. 2009. “The Description–Experience Gap in Risky Choice.” *Trends in Cognitive Sciences* 13(12): 517–23.

Hertwig, Ralph, and Andreas Ortmann. 2001. “Experimental Practices in Economics: A Methodological Challenge for Psychologists?” *Behavioral and Brain Sciences* 24(3): 383–403.

Highhouse, Scott, Christopher D. Nye, Don C. Zhang, and Thaddeus B. Rada. 2017. “Structure of the DOSPERT: Is There Evidence for a General Risk Factor?” *Journal of Behavioral Decision Making* 30(2): 400–06.

Holt, Charles A., and Susan K. Laury. 2002. “Risk Aversion and Incentive Effects.” *American Economic Review* 92(5): 1644–55.

Josef, Anika K., David Richter, Gregory R. Samanez-Larkin, Gert G. Wagner, Ralph Hertwig, and Rui Mata. 2016. “Stability and Change in Risk-Taking Propensity across the Adult Life Span.” *Journal of Personality and Social Psychology* 111(3): 430–50.

Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory—Analysis of Decision under Risk.” *Econometrica* 47(2): 263–92.

Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro. 2008. “Imputing Risk Tolerance from Survey Responses.” *Journal of the American Statistical Association* 103(483): 1028–38.

Knutson, Brian, and Scott A. Huettel. 2015. “The Risk Matrix.” *Current Opinion in Behavioral Sciences* 5(October): 141–46.

Lejuez, Carl W., Jennifer P. Read, Christopher W. Kahler, Jerry B. Richards, Susan E. Ramsey, Gregory L. Stuart, David R. Strong, and Richard A. Brown. 2002. “Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task

- (BART).” *Journal of Experimental Psychology: Applied* 8(2): 75–84.
- Levin, Irwin P., Stephanie S. Hart, Joshua A. Weller, and Lyndsay A. Harshman.** 2007. “Stability of Choices in a Risky Decision-Making Task: A 3-Year Longitudinal Study with Children and Adults.” *Journal of Behavioral Decision Making* 20(3): 241–52.
- Lim, Julian, and David F. Dinges.** 2010. “A Meta-Analysis of the Impact of Short-Term Sleep Deprivation on Cognitive Variables.” *Psychological Bulletin* 136(3): 375–89.
- Lindenberger, Ulman.** 2014. “Human Cognitive Aging: Corriger la fortune?” *Science* 346(6209): 572–78.
- Lönnqvist, Jan-Erik, Markku Verkasalo, Gari Walkowitz, and Philipp C. Wichardt.** 2015. “Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison.” *Journal of Economic Behavior and Organization* 119: 254–66.
- Love, Ross O., and Lindon J. Robison.** 1984. “An Empirical Analysis of the Intertemporal Stability of Risk Preference.” *Southern Journal of Agricultural Economics* 16(1): 159–66.
- Malmendier, Ulrike, and Stefan Nagel.** 2011. “Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?” *Quarterly Journal of Economics* 126(1): 373–416.
- Mamerow, Loreen, Renato Frey, and Rui Mata.** 2016. “Risk Taking across the Life Span: A Comparison of Self-Report and Behavioral Measures of Risk Taking.” *Psychology and Aging* 31(7): 711–23.
- Mata, Rui, Anika K. Josef, and Ralph Hertwig.** 2016. “Propensity for Risk Taking across the Life Span and around the Globe.” *Psychological Science* 27(2): 231–43.
- Mata, Rui, Anika K. Josef, Gregory R. Samanez-Larkin, and Ralph Hertwig.** 2011. “Age Differences in Risky Choice: A Meta-Analysis.” *Annals of the New York Academy of Sciences* 1235(1): 18–29.
- Menkhoff, Lukas, and Sahra Sakha.** 2016. “Determinants of Risk Aversion over Time: Experimental Evidence from Rural Thailand.” Available at SSRN: <http://doi.org/10.2139/ssrn.2785467>.
- Moffitt, Terrie E., Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J. Hancox, HonaLee Harrington, Renate Houts, Richie Poulton, Brent W. Roberts, Stephen Ross, Malcolm R. Sears, W. Murray Thomson, and Avshalom Caspi.** 2011. “A Gradient of Childhood Self-Control Predicts Health, Wealth, and Public Safety.” *PNAS* 108(7): 2693–8.
- Pedroni, Andreas, Renato Frey, Adrian Bruhin, Gilles Dutilh, Ralph Hertwig, and Jörg Rieskamp.** 2017. “The Risk Elicitation Puzzle.” *Nature: Human Behaviour* 1: 803–9.
- Richardson, Michelle, Charles Abraham, and Rod Bond.** 2012. “Psychological Correlates of University Students’ Academic Performance: A Systematic Review and Meta-Analysis.” *Psychological Bulletin* 138(2): 353–87.
- Richter, David, and Jürgen Schupp.** 2015. “The SOEP Innovation Sample (SOEP IS).” *Schmollers Jahrbuch* 135(3): 389–400.
- Roberts, Brent W., and Wendy F. DelVecchio.** 2000. “The Rank-Order Consistency of Personality Traits from Childhood to Old Age: A Quantitative Review of Longitudinal Studies.” *Psychological Bulletin* 126(1): 3–25.
- Roberts, Brent W., Carl Lejuez, Robert F. Krueger, Jessica M. Richards, and Patrick L. Hill.** 2014. “What Is Conscientiousness and How Can It Be Assessed?” *Developmental Psychology* 50(5): 1315–30.
- Rohrer, Julia M.** 2017. “Test–Retest Reliabilities of Scales Included in the Socio-Economic Panel Study.” OSF, Open Science Framework, February 7, <https://osf.io/3ncbt>.
- Sahm, Claudia R.** 2012. “How Much Does Risk Tolerance Change?” *Quarterly Journal of Finance* 2(4): 1250020.
- Schmidt, Frank L., and John Hunter.** 2004. “General Mental Ability in the World of Work: Occupational Attainment and Job Performance.” *Journal of Personality and Social Psychology* 86(1): 162–73.
- Schoemaker, Paul J. H., and John C. Hershey.** 1992. “Utility Measurement: Signal, Noise, and Bias.” *Organizational Behavior and Human Decision Processes* 52(3): 397–424.
- Schonberg, Tom, Craig R. Fox, and Russell A. Poldrack.** 2011. “Mind the Gap: Bridging Economic and Naturalistic Risk-Taking with Cognitive Neuroscience.” *Trends in Cognitive Sciences* 15(1): 11–19.
- Sharma, Leigh, Kristian E. Markon, and Lee Anna Clark.** 2014. “Toward a Theory of Distinct Types of ‘Impulsive’ Behaviors: A Meta-Analysis of Self-Report and Behavioral Measures.” *Psychological Bulletin* 140(2): 374–408.
- Slovic, Paul.** 1964. “Assessment of Risk Taking Behavior.” *Psychological Bulletin* 61(3): 220–33.
- Smidts, Ale.** 1997. “The Relationship Between Risk Attitude and Strength of Preference: A Test of Intrinsic Risk Attitude.” *Management Science* 43(3): 357–70.
- Steinberg, Laurence.** 2013. “The Influence of Neuroscience on US Supreme Court Decisions about Adolescents’ Criminal Culpability.” *Nature Reviews: Neuroscience* 14(7): 513–8.
- Tracy, Jessica L., Richard W. Robins, and Jeffrey W. Sherman.** 2009. “The Practice of Psychological Science: Searching for Cronbach’s Two Streams in Social–Personality Psychology.” *Journal of*

Personality and Social Psychology 96(6): 1206–25.

Viechtbauer, Wolfgang. 2010. “Conducting Meta-Analyses in R with the Metafor Package.” *Journal of Statistical Software* 36(3): 1–48.

Vlaev, Ivo, Nick Chater, and Neil Stewart. 2009. “Dimensionality of Risk Perception: Factors Affecting Consumer Understanding and Evaluation of Financial Risk.” *Journal of Behavioral Finance* 10(3): 158–81.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp. 2007. “The German Socio-Economic Panel Study (SOEP)—Scope, Evolution and Enhancements.” *Schmollers Jahrbuch: Journal of Applied Social Science Studies* 127(1): 139–69.

Wallsten, Thomas S., Timothy J. Pleskac, and Carl W. Lejuez. 2005. “Modeling Behavior in a Clinically Diagnostic Sequential Risk-Taking Task.” *Psychological Review* 112(4): 862–80.

Weber, Elke U., Ann-Renée Blais, and Nancy E. Betz. 2002. “A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors.” *Journal of Behavioral Decision Making* 15(4): 263–90.

Wehrung, Donald A., Kenneth R. MacCrimmon, and Kent M. Brothers. 1984. “Utility Assessment: Domains, Stability, and Equivalence Procedures.” *INFOR: Information Systems and Operational Research* 22(2): 98–115.

Weimann, Joachim, Andreas Knabe, and Ronnie Schöb. 2015. *Measuring Happiness: The Economics of Well-Being*. Cambridge, MA: MIT Press.

White, Jennifer L., Terrie E. Moffitt, Caspi Avshalom, Dawn Jeglum, Douglas J. Needles, Magda Stouthamer-Loeber. 1994. “Measuring Impulsivity and Examining its Relationship to Delinquency.” *Journal of Abnormal Psychology* 103(2): 192–205.

Wölbert, Eva, and Arno Riedl. 2013. “Measuring Time and Risk Preferences: Reliability, Stability, Domain Specificity.” Netspar Discussion Paper no. 07/2013-044. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2354907>.

Wulff, Dirk U., Max Mergenthaler-Canseco, and Ralph Hertwig. 2017. “A Meta-Analytic Review of Two Modes of Learning and the Description–Experience Gap.” *Psychological Bulletin*, 144(2): 140–76.