Received: 28 February 2019 | Revised: 20 December 2019 | Accepted: 20 December 2019
DOI: 10.1002/bdm.2171

RESEARCH ARTICLE

WILEY

Preference or ability: Exploring the relations between risk preference, personality, and cognitive abilities

Philip Millroth | **Peter Juslin** | **Anders Winman** | **Håkan Nilsson** | **Marcus Lindskog**

Department of Psychology, Uppsala University, Uppsala, Sweden

Correspondence
Philip Millroth, Department of Psychology, Uppsala University, P.O. Box 1225, SE-751 42 Uppsala, Sweden.
Email: philip.millroth@psyk.uu.se

Abstract

Key issues in the behavioral sciences are if there exist stable risk preferences that generalize across domains and if these are best measured by revealed risk preference (RRP) in behavioral decision tasks or by surveys eliciting stated risk preference (SRP). We applied network analysis to data from a representative Swedish sample to investigate the relations between RRP, SRP, personality characteristics, and cognitive abilities, using in total over 70 measurements. The results showed that different measures of RRP were poorly intercorrelated and formed a community together with measures of numerical and cognitive abilities. Measures of SRP were weakly correlated with measures of RRP and identified in a distinctly separate community, along with personality characteristics and gender. The ensuing analyses provided support for a model suggesting that RRPs are contaminated by demands on numerical and cognitive abilities. RRPs may thus suffer from poor construct validity, whereas SRPs may better capture people's everyday risk preferences because they are related to more stable traits.

KEYWORDS

cognitive abilities, explorative, personality, risk preferences

1 | INTRODUCTION

The notion of stable risk preferences lay at the core of normative and descriptive theories of decision making under risk (e.g., von Neumann & Morgenstern, 1944/2007; Tversky & Kahneman, 1992; Mishra, Barclay, & Sparks, 2017). The impact of these theories in the social sciences is not surprising, given the expectation that they can be used both to predict people's behavior and to allow people to make more informed decisions in different domains.

Typically, risk preferences are captured by either measuring stated or revealed risk preferences (SRPs or RRPs). SRPs are often captured using self-report measures (e.g., "How often do you use a seatbelt when driving?" and "How willing are you to take risks"), whereas RRPs are captured using behavioral measures such as evaluating or making choices between monetary lotteries (e.g., "How much are you willing to pay to participate in a lottery with a .8 probability of $100 and a .2 probability of obtaining $0?" or "Would you prefer A: receiving $50 for sure, or B: playing a lottery which yields nothing with a .50 probability and $100 with a.50 probability") and playing card games (e.g., betting on whether the next drawn card would be of a higher or lower rank than the present card).

In the past decade, the two approaches have been developed in parallel with the number of possible measures assessing risk preferences through SRPs and RRPs steadily growing, often with the (at least) implicit assumption that they connect to the same underlying construct of a stable risk preference. Prominent examples of the SRP approach include the Domain-Specific Risk-taking (Weber, Blais, & Betz, 2002) and the Risk Propensity Scale (Nicholson, Fenton-O'Creevy, Soane, & Willman, 2002; Nicholson, Soane, Fenton-O'Creevy, & Willman, 2005). Although the evaluation of or choice between monetary lotteries has been the most used RRP elicitation

The research was sponsored by the Swedish Research Council. The study was not formally preregistered. The dataset with de-identified data is available upon request. The materials used in these studies are widely available (see Supporting Information). The author(s) declared no conflicts of interest with respect to the authorship or the publication of this article.

tasks, other prominent approaches have used other tasks, such as card games (the Iowa Gambling Task: Bechara, Damasio, Damasio, & Anderson, 1994), choosing between boxes that hide possible rewards (the Cups Task: Levin, Weller, Pederson, & Harshman, 2007) and choosing whether or not to further inflate a balloon (the Balloon Analogue Risk Task: Lejuez et al., 2002).

However, recent years have seen an intense effort to uncover the convergent validity and temporal stability of SRPs and RRPs and also map how they relate to other psychological constructs, such as cognitive ability[1] and personality (e.g., Big-5 and impulsiveness). In the remainder of Section 1, we briefly review the current state of affairs and argue that there is a need for (a) conceptual replications of findings and (b) studies that map how risk preference, as measured by SRPs and RRPs, relates to many different types of measures of the personality construct and measures of cognitive abilities. Thereafter, we present exploratory analyses of a dataset with over 200 participants representative of the Swedish population that includes a large number of dependent variables (DVs)[2]: 13 DVs measuring RRPs and SRPs and 60 DVs measuring cognitive abilities, personality characteristics, and demographics. The nature of the dataset makes it well-suited to test a number of questions regarding the nature of risk preferences (further detailed in Section 1.2, "Aims, scope, and overview of the present article") while avoiding sampling bias of our stimuli by taking an exploratory approach.

1.1 | SRPs and RRPs: Relation to each other and to other psychological constructs

Although some studies have reported that SRPs and RRPs are related (e.g., Dohmen et al., 2011; Mishra & Lamuiere, 2011; Agren, Millroth, Andersson, Ridzén, & Björkstrand, 2019) and that single-item measures from both methods can be combined to form a unified risk preference (Falk et al., 2018), others have found that SRPs and RRPs are not related or only weakly so (Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017; Lönnqvist, Verkasalo, Walkowitz, & Wichardt, 2015).

Important differences between SRPs and RRPs have been documented, and both types of measures have been linked to other psychological constructs. First, SRPs seem to show both higher test–retest stability and convergent validity relative to RRPs (Frey et al., 2017; Mata, Frey, Richter, Schupp, & Hertwig, 2018; Pedroni et al., 2017). Second, there exists strong evidence for a relation between SRPs and personality measures (e.g., Big-5 traits; Frey et al., 2017; Mata et al., 2018; Nicholson et al., 2005), and although some studies have documented a relation between RRPs and personality measures such as impulsiveness (Demaree, DeDonno, Burns, & Everhart, 2008; Demaree, DeDonno, Burns, Feldman, & Everhart, 2009; Rosenbaum & Hartley, 2018), others have failed to identify such a link (Frey et al., 2017) or only found low degrees of association (Becker, Deckers, Dohmen, Falk, & Kosse, 2012).

Third, although age is seemingly related to both SRPs and RRPs (with reduced risk-taking with increasing age), the correlations between measures are small and age differences in behavioral paradigms seem to emerge as a function of specific task characteristics, such as learning and computational demands (Mamerow, Frey, & Mata, 2016; Mata, Josef, Samanez-Larkin, & Hertwig, 2011). Age is also related to cognitive abilities (following an inverse U-shape, see, e.g., Li et al., 2015) and impulsiveness (increasing in adolescence but then declining with age: Rosenbaum & Hartley, 2018). Moreover, impulsiveness is seemingly directly related to executive functions (cognitive processes involved in the control of thought and action, see Leshem & Glicksohn, 2007), hence further complicating the possible causal pathways between age, impulsiveness, and risk attitudes.

Fourth, gender differences have been most pronounced for SRPs (for a review, see Lilleholt, 2019). Fifth, and finally, SRPs and RRPs seemingly differ on how they relate to cognitive abilities: When RRPs are used, there is strong evidence that cognitive ability relate negatively to risk aversion (i.e., people with higher levels of cognitive ability tend to be less risk averse: see Lilleholt, 2019 for a comprehensive meta-analysis), and one cognitive ability that seem to be especially predictive is numeracy (i.e., the ability to reason and to apply numerical concepts—see, e.g., Cokely et al., 2018). In contrast, the relation between cognitive abilities and SRPs remains dubious. On the one hand, Frey et al. (2017) did not find any relation between measures of cognitive ability (more specifically, working memory and numeracy) and five different SRP scales. On the other hand, cognitive abilities—when measured as general intelligence—have been found to correlate with SRPs when single-item measures have been used (i.e., responding on a Likert scale to a general question about one's willingness to take risks: Dohmen et al., 2010; Beauchamp, Cesarini, & Johannesson, 2017; Dohmen, Falk, Huffman, & Sunde, 2018).

In summary, previous research has documented different relations between RRP and SRP measures of risk attitude, on the one hand, and measures of personality and cognitive and numerical ability, on the other hand, suggesting a difference between the two sorts of measures of risk attitude. However, previous studies have not entered a comprehensive set of measures of all four types of measures (RRP, SRP, personality, and cognitive and numerical abilities) into a common exploratory analysis to investigate how these variables spontaneously organize into common clusters. As proposed below, this approach offers complementary benefits.

1.2 | Aims, scope, and overview of the present article

Although studying constructs in isolation is not problematic per se, it can lead to a lack of cumulative knowledge in a given field

[1]Henceforth defined as "individual differences in the capacity to successfully performs tasks that require the manipulation, retrieval, evaluation, or processing of mental information," see Lilleholt (2019).

[2]It is easy to confuse terms such as *measures*, *variables*, *items*, and *submeasures*. When we discuss measures, we mean a certain test or scale. For example, Big-5 is considered a *personality measure* and Raven's matrices, a *measure of general cognitive ability*. A measure may be organized into submeasures, as when Big-5 is organized into submeasures of conscientiousness, openness, extraversion, agreeableness, and neuroticism. Any measure, as well as submeasure, consist of a number of items: a specific stimuli presented to the responder and for which a response is noted. Responses to items are used to define measures, typically by aggregating the responses. A measure becomes a dependent or independent variable when it enters into an analysis where it is assumed that the measures have the logical status of independent or dependent variables.

(e.g., Eisenberg et al., 2019). Integrative explorative approaches that simultaneously test for relations between risk measures, personality characteristics, and cognitive abilities are unfortunately scarce; typically either personality characteristics or cognitive abilities are studied in relation to either SRPs or RRPs, and in the more integrative attempts that aim to cover both RRPs and SRPs (e.g., Chapman, Dean, Ortoleva, Snowberg, & Camerer, 2018; Falk et al., 2018; Frey et al., 2017; Mamerow et al., 2016), only a few measures for personality and cognitive abilities are used. This study complements the literature by incorporating DVs from four types of measures (SRP, RRP, personality, and cognitive abilities) into a single explorative network analysis.

Although overlap with past work is desirable in order for the field to eventually arrive at reliable estimates of convergent validity of measures and constructs, the introduction of additional measures (e.g., numerical cognition, metacognitive ability, and subjective well-being) can ensure a comprehensive spread of relevant psychological aspects. It can also provide information on possible mediators and moderators; in order to rule out back-door pathways of causality, it is often necessary to provide measurements of all relevant concepts (Pearl, 2000). Although it is not the aim of this study to provide an overarching causal account of how risk preferences are formed, future studies in that direction should benefit from analysis of a dataset such as the one we present here. Moreover, fitting all measures into one study also allows for a direct test of whether some constructs are more related to each other compared with other constructs. For example, it may be that both SRPs and RRPs relate to cognitive ability, but RRPs much more so. Alternatively, it may be that even though SRPs and RRPs are related, they actually show stronger relation to other constructs than to each other.

We employed network analysis in the initial exploratory analysis because in contrast to other techniques, such as factor analysis, network analysis makes no commitment to causal assumptions.[3] As has been argued elsewhere (Behrens, 1997; Judd & Kenny, 2010; Kelder, Conklin, Evelo, & Pico, 2010; Tukey, 1980), any research field need, from time to time, to complement the confirmatory approach with an exploratory approach. Importantly, an exploratory approach can help us avoid sampling bias (Fiedler, 2011; Koch, Imhoff, Dotsch, Unkelbach, & Alves, 2016), meaning that in order to evaluate the support for a given hypothesis, we need to evaluate it in comparison with competing hypotheses that are simultaneously tested.

Specifically, we present exploratory network analysis of a dataset, with over 200 participants representative of the Swedish population that includes a large number of DVs: 13 DVs measuring RRPs and SRPs and 60 DVs measuring cognitive abilities, personality characteristics, and demographics. Representative population samples, in contrast to college student samples that are typically used in psychological research, offer the beneficial property of increased variance in regard to not only age but often also performance on the measures included. In the first step of the analysis, the network analysis, we enter DVs for all submeasures separately allowing them to organize freely and spontaneously into natural communities (clusters) of related variables. For example, we use six DVs of the Risk Propensity Scale instead of aggregating scores from the six DVs into a DV measuring overall risk propensity. It was a deliberate choice not to collapse DVs from submeasures to an aggregate DV (e.g., collapsing the six measures of risk propensity), in order to not constrain the possible relations a priori (e.g., relations between the submeasures of the SRP and other variables cannot be detected if they were only analyzed with collapsed scores).

Foreshadowing the results, the network analysis suggested that the RRP measures formed a community with the cognitive abilities that was distinctly separate from the SRP measures, which were related to personality trait measures. This fits badly with the default assumption in the field that both kinds of measures should tap into a common underlying trait-like preference for risk (e.g., Dohmen et al., 2011; Stigler & Becker, 1977). Based on this exploratory analysis and the growing literature on the relations between numeracy and performance in decision tasks (for a review, see Cokely et al., 2018), we formulated a contamination hypothesis: RRP measures that draw on monetary lotteries are contaminated by their demands on numerical and cognitive abilities, making them disclose poor convergent validity and limiting their validity as predictors of people's real-life risk preferences.

In the subsequent analyses, we turned to more confirmatory analyses, where we quantified the extent to which the data provided evidence for this contamination hypothesis in comparison with two plausible alternative hypotheses. At this stage, the DVs are created by aggregating over submeasures because the network analysis showed that such an approach would not be problematic (e.g., social risk-taking and financial risk-taking did not relate to different types of measures). The risk-trait hypothesis claims that although the measures of RRP and SRP may be related also to specific cognitive abilities and personality characteristics, they should both tap into a common stable underlying construct for risk preference (e.g., Dohmen et al., 2011; Stigler & Becker, 1977). Another possibility is that the intercorrelations are driven by common methods, where some DVs are based on self-reports and other DVs are based on behavioral observations. The common method hypothesis thus evaluates an alternative artefactual explanation emphasizing method bias (for a review on common method bias, see Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), assuming one underlying factor common to all self-report DVs and one factor common to all of the behavioral DVs.

2 | METHOD

2.1 | Participants

A random sample of 2,000 inhabitants of Uppsala, with the criteria of obtaining an equal gender distribution, participants being between

[3]An additional benefit for our purposes is that it is less sensitive to the ratio of observations per DV—see Section 2 for a discussion.

TABLE 1 Measures of risk preferences, detailing the type of test (behavioral or self-report), name of the test, aim of the test, and how the measure is abbreviated in the figures

Type of test	Test name	Aimed to measure	Abbreviation
Behavioral	Cups Task	Proportion of risky choices	RRP 1
Behavioral	Discounting	Intertemporal risk attitude	RRP 2
Behavioral	Iowa Gambling Task	Adaptive risk-taking	RRP 3
Behavioral	Loss Aversion Task	Attitudes for negative risky prospects	RRP 4
Behavioral	WTP Task	Attitude to risky prospects	RRP 5
Behavioral	Wakker Test	Utility function for risky prospect	RRP 6
Behavioral	Wakker Test	Probability function for risky prospects	RRP 7
Self-report	Risk Propensity Scale	Recreational risk	SRP 1
Self-report	Risk Propensity Scale	Health risk	SRP 2
Self-report	Risk Propensity Scale	Career risk	SRP 3
Self-report	Risk Propensity Scale	Financial risk	SRP 4
Self-report	Risk Propensity Scale	Safety risk	SRP 5
Self-report	Risk Propensity Scale	Social risk	SRP 6

Abbreviations: RRP, revealed risk preference; SRP. stated risk preference.

20 and 60 years, and living a maximum of 20 km from the town center, was ordered from Statistics Sweden (a government agency), with the goal of obtaining a sample representable of the general population of Sweden. Because Uppsala is a university town, student-dominated areas were excluded. The individuals were contacted by post. Three hundred and thirty-two responded to the letters. Out of these, 213 participated in the study sessions (age: M = 39, SD = 12). High attrition rates can be problematic if its self-selection leads to the sample no longer being representable of the intended population (e.g., Zhou & Fishbach, 2016). Sixty-two percent were women and 38% were men. Self-selection resulted in an overrepresentation of women (in Section 4, we argue that this did not affect the results). The attrition rate did not seem to affect other aspects of the sample, it being otherwise representative of the Swedish population (population statistics from 2012 are valid for individuals between 16 and 65 years and were obtained from Statistics Sweden: Thirty-four percent, 40%, and 21% reported elementary school, high school, and university, respectively, as highest level of education (corresponding percentages in the total Swedish population are 35.3%, 41.3%, and 23.4%, respectively, for elementary school, high school, and university). Eighty-six percent were born in Sweden (corresponding percentages in the total Swedish population is 81%). The median monthly income was 25,000 SEK (corresponding number in the total Swedish population is 25,400 SEK). For compensation, participants could choose between a gift certificate with a value of 1,000 SEK or donating the same amount to an optional charity.

2.2 | Measures and DVs

Thirteen DVs were collected in order to measure RRPs and SRPs (see Table 1); 56 DVs were collected in order to measure cognitive abilities and personality characteristics (Table 2); and four DVs were collected in order to measure demographics (highest level of education, gender, monthly income, and age). None of the behavioral tasks were incentivized.

Frey et al. (2017) used an impressive set of almost 40 DVs of risk attitude. Because our data collection also aimed at researching other topics (a subset of the data has been used for other publications on the approximate number system and numeracy: Winman, Juslin, Lindskog, Nilsson, & Kerimi, 2014; Lindskog, Kerimi, Winman, & Juslin, 2015), we focused on key measures in the literature that have been used to make inferences about peoples' risk preferences (e.g., they are all included in influential reviews on the topic, see Glimcher & Fehr, 2014; Camerer, Loewenstein, & Rabin, 2011; Keren & Wu, 2015). We do not claim to have an exhaustive list of DVs that cover all possible measures of cognitive abilities and personality characteristics. However, our dataset included measures covering many of the most frequently used measures in the literature.

2.3 | Statistical analyses

2.3.1 | Network analysis

Network analysis was used in the initial exploratory analysis for two reasons: First, network analysis does not make the assumption of causality that the observed factors are due to an underlying latent variable, it simply groups variables together on the basis of common relations. The latent variable approach, as characterized by factor analysis, has been shown to be problematic in research on personality and psychopathology because it assumes causality where there is none (Epskamp, Rhemtulla, & Borsboom, 2017).[4] Second, network

[4]It has also been empirically shown that network analysis outperforms exploratory and confirmatory algebraic analysis tools for producing plausible modularity at item level for the five-factor model of personality traits and can also better identify the key roles of individual items and clusters (Goekoop, Goekoop, & Scholte, 2012), but see Bringmann and Eronen (2018) for a discussion.

TABLE 2 Measures of cognitive abilities and personality, detailing the type of test (behavioral or self-report), name of the test, aim of the test, and how the measure is abbreviated in the figures

Type of test	Test name	Aimed to measure	Abbreviation
Behavioral	ANS Test	Acuity of mental number line	ANS
Behavioral	ANT	Objective numeracy	ANT
Behavioral	Base Rate Test	Probabilistic reasoning	BRT
Behavioral	Inspection Time Test	Mental and perceptual speed	IIT
Behavioral	Lipkus	Numeracy	NL
Behavioral	LTM Test	Recognition memory	LTM 1
Behavioral	LTM Test	Recall memory	LTM 2
Behavioral	Mouselab	Information search	IS
Behavioral	Raven's Matrices	IQ	IQ
Behavioral	Syllogism Test	Reasoning confidence	ST
Behavioral	WM Test	Working memory capacity	WMT
Behavioral	Expected Value Task	Knowledge about EV	EVT
Behavioral	Transitivity Test	Transitive ability	TT
Behavioral	Wechsler Memory Scale	Verbal memory	WMS
Behavioral	Conjunction Fallacy Test	Probabilistic estimation	CF 1
Behavioral	Conjunction Fallacy Test	Probabilistic decision	CF 2
Behavioral	Metacognitive Ability (over/under)	Over/under confidence	MT 1
Behavioral	Metacognitive Ability	Confidence discrimination	MT 2
Behavioral	Metacognitive Ability	Confidence linearity	MT 3
Behavioral	Metacognitive Ability	Over/under estimation	MT 4
Behavioral	Metacognitive Ability	Assessed numerical performance	MT 5
Behavioral	Metacognitive Ability	Assessed non-numerical performance	MT 6
Self-report	SNS	Subjective numeracy	SNS
Self-report	Life Orientation Test	Pessimistic and optimistic attitudes	LOT
Self-report	Self-Esteem Test	Self-esteem	SE
Self-report	SWB Test	Subjective well-being	SWB
Self-report	SC Test	Social network	SN
Self-report	PS Test	Perceived stress	PS
Self-report	Big-5 Test	Conscientiousness	BIG 1
Self-report	Big-5 Test	Stability	BIG 2
Self-report	Big-5 Test	Agreeableness	BIG 3
Self-report	Big-5 Test	Openness	BIG 4
Self-report	Big-5 Test	Extraversion	BIG 5
Self-report	Rational Experiential Inventory	Experiential ability	REI 1
Self-report	Rational Experiential Inventory	Rational favorability	REI 2
Self-report	Rational Experiential Inventory	Rational ability	REI 3
Self-report	Rational Experiential Inventory	Experiential favorability	REI 4
Self-report	Baratt Impulsiveness Scale	Second-order attention	BIS 1
Self-report	Baratt Impulsiveness Scale	Second-order motor	BIS 2
Self-report	Baratt Impulsiveness Scale	Second-order planning	BIS 3
Self-report	Baratt Impulsiveness Scale	First-order attention	BIS 4
Self-report	Baratt Impulsiveness Scale	First-order motor	BIS 5
Self-report	Baratt Impulsiveness Scale	First-order self-control	BIS 6
Self-report	Baratt Impulsiveness Scale	First-order cognitive complexity	BIS 7
Self-report	Baratt Impulsiveness Scale	First-order perseverance	BIS 8

(Continues)

TABLE 2 (Continued)

Type of test	Test name	Aimed to measure	Abbreviation
Self-report	Baratt Impulsiveness Scale	First-order cognitive instability	BIS 9
Self-report	Cognitive Failure Questionnaire	Perceived cognitive ability	CFQ

analyses have been found to be less sensitive to sample size compared with factor analysis (Golino & Epskamp, 2017).

Network analysis characterizes networks in terms of nodes (people, organizations, etc.) and the edges (relationships or interactions) that connect them. It also allows for the detection of communities: A network is said to have a community structure if the nodes of the network can be easily grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally. The more general definition is based on the principle of modularity that pairs of nodes are more likely to be connected if they are both members of the same community (ies) and less likely to be connected if they do not share communities. Calculating the modularity of the community structure is used to assert that the identified structure is not due to chance. Modularity reflects the concentration of edges within modules compared with a random distribution of links between all nodes regardless of modules, and it measure Q ranges between −.50 and 1.0. Values over .30 are considered good (Newman, 2006). Subsequent work has confirmed that Q-values over.20 can be considered to be statistically robust (e.g., Poisot, 2013).

A Bayesian analysis was first performed to identify a reliable correlation matrix for the network analysis. All Bayesian correlation analyses were conducted with default settings in the statistical software JASP v.0.8.4.0 (JASP Team, 2018). Kendall's tau coefficient was used for all correlations. Consistency has been found to be higher among rank-based measures for risk preferences (Pedroni et al., 2017), and given that the nature of possible relations are a priori unknown here, Kendall's tau coefficient offers many advantages over Pearson's r coefficient (see van Doorn, Ly, Marsman, & Wagenmakers, 2018).

Correlations with a Bayes factor over 10 (considered strong evidence) were included in the matrix. Network analysis was conducted using Gephi 0.9.2 (Bastian, Heymann, & Jacomy, 2009). The Gephi data file containing all network information is available as Supporting Information online. The correlation matrix (available as Supporting Information) was used as input for the optimization algorithm used for community detection (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) and for the algorithm used to create the spatial output (Jacomy, Heymann, & Bastian, 2014).

We acknowledge that the reliance on correlations supported with a BF > 10 increases the rate of Type 2 error, but we are at present more concerned with avoiding Type 1 errors. The weakest detected correlation with a BF > 10 was.141, suggesting Type 2 errors could be a factor for weaker correlations. However, complementary robustness analyses where we tested if either a lower threshold (BF > 3) or a different correlation statistic (Pearson's r) affected the community structure derived from the network analysis showed (see Appendix A) showed that the results presented in the main part of this article were robust to those changes.

2.3.2 | Structural equation modeling

To quantify the evidence in favor of the conclusions suggested by network analysis, we compared a number of models using confirmatory factor analysis (CFA). The CFA was implemented in the R-package "lavaan." The standardized data, z-scores, were used as input. There were a total of 40 missing data patterns.[5] The Little missing completely at random test showed that data could be considered missing completely at random (Little, 1988). In lavaan, maximum likelihood estimation was used for the missing data patterns, and robust maximum likelihood estimation was used for parameter estimation. Importantly, it has been shown that the maximum likelihood estimation approach renders reliable parameter estimates over a wide range of conditions, offering an advantage over weighted least squares with sample sizes over 200 (Li, 2016). At the point of the CFA analyses, the numbers of DVs had been reduced to 37 (see Section 3 for details), resulting in an observations-per-DV ratio of 5.76. The number of variables within each proposed factor ranged between 7 and 19.

3 | RESULTS

Section 3 is provided in three sections. First, descriptive statistics for the effect sizes are presented in order to offer a general overview of the data. Second, the results from the network analysis are presented. Third, the results from ensuing confirmatory factor analysis are presented.

3.1 | Descriptive statistics for effect sizes

Table 3 provides measures of central tendency and variability for the correlations between the four types of measures of main concern in this study (i.e., ability, personality, RRP, and SRP), where Kendall's tau has been converted to Pearson's r for illustrative purposes (see Walker, 2003). As in similar large-scale data collections (Frey et al., 2017; Griffin, Guilette, & Healy, 2015; Pedroni et al., 2017), the magnitude of the correlations observed between the DVs included is very modest. For example, the median correlation between measures of SRP and RRP is.085.

The weak correlations may be caused not only by the absence of real relations but also by the possibility that the measures collected by

[5]A dataset with variables $Y_1, Y_2, \ldots, Y_p$ (in that order) is said to have a missing pattern when the event that a Y_j variable is missing for a particular individual implies that all subsequent variables Y_k, $k > j$, are missing for that individual, or when a variable is observed for a particular individual, it is assumed that all previous variables Y_k, $k > j$, are also observed for that individual.

TABLE 3 Descriptive statistics for summarized effect sizes (Pearson's *r*—transformed from Kendall's tau) based on all the initial measures and independent of evidence strength

Relationship	*M*	*MD*	*SD*	Tenth percentile	Ninetieth percentile	Min	Max
Personality–Ability	.079	.064	.069	.011	.174	.000	.407
RRP–Personality	.076	.063	.059	.013	.166	.000	.322
SRP–Ability	.067	.052	.054	.008	.139	.000	.291
SRP–Personality	.116	.093	.091	.019	.238	.002	.413
RRP–Ability	.104	.070	.093	.016	.262	.002	.385
SRP–RRP	.099	.085	.063	.017	.180	.003	.246

Abbreviations: RRP, revealed risk preference; SRP, stated risk preference.

psychologists sometimes include more noise than signal (Sripada, Kessler, & Jonides, 2016). Table 4 reports the proportion of correlations between DVs from the four types of measures that are pairwise compared in Table 3 for which there exists reliable empirical evidence for a correlation, as compared with the null hypothesis of no correlation (i.e., correlations with BFs over 10). It is clear from Table 4 that there are only for three pairwise comparisons that there exists a nontrivial proportion of statistically reliable correlations (8% of empirically supported correlations between personality and ability, 10% of empirically supported correlations between SRP and personality, and 11% of empirically supported correlations between RRP and ability). For the other comparisons, there is empirical support for less than 2% of the intercorrelations.

The rest of Section 3, which presents the results from the exploratory network analysis and the structural equation modeling, is devoted to testing if the hypotheses outlined in Section 1 can be distinguished when we consider those correlations for which there is empirical support (see Appendix A for robustness analyses showing that the community structure derived from the network analysis is the same also when allowing for a lower threshold (BF < 3) and different a correlation coefficient (Pearson's *r*).

3.2 | Network analysis

As illustrated in Figure 1, three communities were detected (modularity of $Q = .404$): (a) personality measures capturing impulsiveness and individual well-being, (b) SRPs together with a group of personality measures, and (c) cognitive abilities together with RRPs.[6,7] That the RRPs were all detected in the community where the measures of cognitive abilities were also included signals that RRPs are not only related to cognitive abilities but much more so than to SRPs or to personality characteristics.

The results of the network analysis in Figure 1 are not very suggestive of the risk-trait hypothesis (because no community capturing risk preference is shared by the RRP and SRP) and do not accord with the common method hypothesis (because not all self-reports are situated in one community with all behavioral measures situated in another community). That RRP is part of a community shared by cognitive abilities distinct from the community with SRP, personality characteristics, and gender is, however, consistent with a contamination hypothesis suggesting that RRP primarily taxes cognitive abilities rather than risk preference. The detected correlations between the RRPs and cognitive abilities suggest that the lack of correlation between RRPs and SRPs is not the mere result of a low reliability of the RRPs.

The submeasures with generally considered good psychometric properties (REI, BIS, and RPS) cluster together, as is also true for submeasures of “in-the-lab” measures (LTM and CF).

We thus conclude that network analysis provides support for the validity of the traditional psychometric measures included and that the pattern of relations provide tentative support for the contamination hypothesis, rather than for the risk-trait and common-method hypotheses. In the following, we aim at validating these impressions with confirmatory analyses.

TABLE 4 Summary of number of realized connections over the number of possible connections (proportions in parentheses) across domains

Connected domains	Possible connections and proportions with BF over 10
Personality–Ability	459 (.081)
RRP–Personality	189 (.005)
SRP–Ability	102 (.019)
SRP–Personality	162 (.105)
RRP–Ability	119 (.109)
SRP–RRP	42 (.000)

Abbreviations: RRP, revealed risk preference; SRP, stated risk preference.

3.3 | Confirmatory factor analysis

Three CFA models were evaluated. The risk-trait model assumed four factors, one underlying factor common to RRPs and the SRPs that capture risk preference and the three other factors suggested by

[6]The network density was .192, meaning that of all possible connections, 19.2 per cent were realized. As network density is only interpretable in relation to comparable network, this figure is not meaningful here, but it may be of useful comparison for future studies on the topic.

[7]The metacognitive tests (MT 1 and MT 7) were not related to any other variables than themselves and were thus dropped from the network.

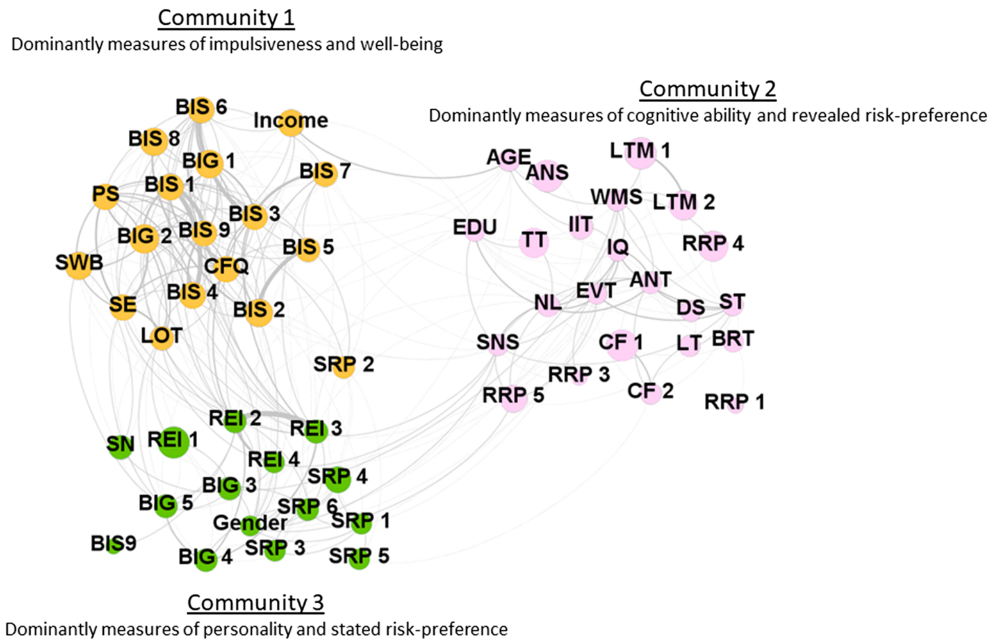

FIGURE 1 Visualization of the relational network where lines constitute links that are constituted by correlations (Tau correlations with BF > 10) and nodes (dependent variables [DVs]) are constituted by dots. Size of lines increases dependent on the strength of the correlation. Node size is dependent on how many links are connected to the node: Increased number of connections results in an increased node size. Colors mark the three detected communities (Pink = DVs from measures of revealed risk preference and cognitive ability; Green = dominantly DVs from measures of stated risk preference and personality; Orange = dominantly DVs from measures of impulsivity and well-being). A complete list of abbreviations is provided in Tables 1 and 2. DVs not evident in the figure did not correlate with other DVs (e.g., "RRP 1", "RRP 6", and "RRP 7") [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 1 (i.e., for impulsiveness and well-being, for personality measures, and for cognitive abilities). The contamination model assumed three factors, one where RRP was aligned with the cognitive abilities, one where SRP was aligned with the personality measures, and one remaining factor for impulsiveness and well-being. The common-method model assumed two factors, one underlying factor common to all of self-report measures and one factor common to all behavioral measures.

Aggregating submeasures resulted in 40 DVs used for the CFA analyses (number of observations-to-variable ratio = 5.68), see Table B1 in Appendix B. See Section 2 for the specifics of the implementation.

The results of the CFA (see Table 5) show that the two fit indices, the root mean square error of approximation (RMSEA, derived from the examining discrepancy between the hypothesized model, with optimally chosen parameter estimates, and the population covariance matrix) and comparative fit index (CFI, derived from examining the discrepancy between the data and the hypothesized model while adjusting for the issues of sample size inherent in the chi-squared test of model fit), are not in agreement: The RMSEA for both the contamination model and the risk-trait model are on, or below, the typical cutoff values for what is considered adequate fit (.07 to .08, see Hooper, Coughlan, & Mullen, 2008), but the CFI is well below what is typically considered adequate fit (.90 to.95, see Hooper et al., 2008).

That the CFI is low is not very surprising, because the median correlations in the present dataset are low (Table 3), that the models involve a large number of parameters, and that previous efforts to capture the "risk preferences puzzle" through factor analysis have left similarly large proportions of the variance unexplained (e.g., Frey et al., 2017). Relative fit measures can be expected to be low for other reasons as well: They are more suitable for exploratory analyses (Rigdon, 1996), and they are heavily dependent on sample size and model complexity (Brown & Cudeck, 1992; Schermelleh-Engel, Moosbrugger, & Müller, 2003). It is out of the scope here to fully discuss situations when the CFI and RMSEA are in disagreement and why the RMSEA may be a more appropriate measure for our purposes (for that purpose, the reader is referred to Lai & Green, 2016), but in short, it can be argued that the RMSEA is better apt at handling the complexity of the present models.

Although acknowledging that an RMSEA of.07 is not a really good fit of data, we nevertheless argue it is sufficiently low to make model comparisons fruitful, and we thus turn to determining which of the models are most probable. The contamination model produced a lower Bayesian information criterion than the risk-trait model: a difference of 13. According to Raftery (1995), a Bayesian information criterion difference of 10 is equivalent of a Bayes factor of 150. Thus, the contamination model is best supported by the data among the models tested here.

4 | DISCUSSION

In this study, we investigated people's risk preferences using exploratory network analysis. This allowed us to explore how the submeasures of a number of traditional psychometric measures spontaneously organize themselves and how measures of cognitive abilities and personality spontaneously align themselves with different measures of people's risk preferences. In the following, we (a) briefly summarize the results, (b) relate our findings to previous research, (c) discuss limitations of this study and the field in general, and (d) state the conclusions with its possible implications.

4.1 | Summary of results

The results of the network analysis showed that SRPs were related to personality measures. The results from network analysis also showed that there was overwhelmingly low convergent validity for RRPs; the measures of the behavioral risk preferences were all unrelated. The results also showed relations between cognitive abilities and RRPs and that this link was stronger than both the link between RRPs and SRPs and the link between RRPs and personality characteristics. We used the data to differentiate between three competing hypotheses: (a) the contamination hypothesis, holding that RRPs that draw on monetary lotteries are contaminated by their demands on numerical and cognitive abilities, (b) the risk-trait hypothesis, holding that RRPs and SRPs should both tap into a common underlying construct for risk preference, and (c) the common method hypothesis, holding that any observed differences between SRPs and RRPs stem from the fact that different types of measures are used (self-report measures vs. behavioral measures). The results of the network analysis and the CFA, along with the descriptive statistics, supported the contamination hypothesis.

4.2 | How the results align with and complement previous research

Many of our results align with previously robust findings: (a) that RRPs are related to measures of cognitive ability (Lilleholt, 2019), (b) that SRPs are closely related to measures of personality (e.g., Frey et al., 2017; Nicholson et al., 2005), (c) and that gender is related foremost to SRPs (for a review, see Lilleholt, 2019). The results also replicate findings that have been recently observed but given the replication crisis have yet to be affirmed as robust: (a) that SRPs and RRPs are at best weakly related (Frey et al., 2017) and (b) that there is low convergent validity for RRPs (Pedroni et al., 2017). By including a comprehensive set of all four types of measures, we also complement previous research by showing that RRPs are not only related to cognitive abilities, it is more so than to any other included construct (both in regard to number of connections and the strength of these connections).

The results also shed some light on issues that have been debated. First, the results provide additional evidence that SRPs are at best weakly related to cognitive abilities, as previously shown by Frey et al. (2017) but in contrast to the results of Dohmen et al (2010). It has also been debated whether impulsiveness and age has a direct effect on risk preferences, or whether their influence is indirect, through cognitive abilities (see, e.g., Mata et al., 2011; Rosenbaum & Hartley, 2018). On basis of the location in the network, our results support the notion that the influence of impulsiveness and age is indirect, hence supporting a view that specific task characteristics, such as learning and computational demands, will interact with these variables to produce an effect on the behaviorally elicited risk preference.

4.3 | Limitations of this study and issues for future research

One limitation of this study is statistical power. Even though the data provide sufficient power to derive a stable mapping of communities using network analysis, our sample size is just short of what is typically required for performing CFA with reliable results. Accordingly, the results of the CFA are best interpreted in the light of the results from the network analysis.

A concern is of course if the results are stable across key demographic variables, gender and age, that have connections with risk preferences. In the current study, we did not have the statistical power to test separate models based on gender and age. However, to get an indication of the possible influence of these variables, we ran multigroup CFA with gender (male/female) and age (young adults <35 years; old adults >35 years) as group. In these analyses, model parameters were constrained to be the same for both groups. For both gender and age as group, these analyses indicated slightly poorer model fit. However, the relative fit of the contamination and risk-trait model remained the same. Although these analyses do not speak directly to gender and age effects in our data, we take them to indicate that our overall conclusion from the CFA holds even when taking gender and age into account. Future research using a similar approach should take into account that a large number of observations are needed to draw solid conclusions about effects of gender and age.

Similarly, a possible corollary of the contamination hypothesis is that self-report and behavioral measures of risk preference show high (er) correlations for individuals with high cognitive ability because, for them, the cognitive demands of behavioral tasks do not mask their underlying risk preference. This intriguing possibility also needs to be examined using a dataset with more statistical power. An additional corollary (but not mutually exclusive) concerns measurement error. One would assume that low levels of cognitive abilities would lead to an increase in measurement error; if participants do not feel confident in solving the problem at hand, it makes sense that they would exhibit more variation in attention and focus. Had we included duplicate elicitations of our measures, we could have pinpointed the role of measurement error in more detail (see, e.g., Gillen, Snowberg, & Yariv, 2019).

TABLE 5 Results of CFA for each of the five tested models, with robust approximations of RMSEA, CFI, and BIC. Values within parentheses illustrate 90% confidence intervals

Model	Number of factors	RMSEA	CFI	BIC
Risk trait	4	.070 (.065; .075)	.632	20,923
Contamination	3	.069 (.064; .074)	.632	20,910
Common method	2	.074 (.069; .079)	.581	20,994

Note. The same rank order of models, in regard of RMSEA and BIC, was withstanding also when the demographics variables (gender, education, income, and age) were included in the analysis (placed in the factor proposed by the community structure of the network analysis).
Abbreviations: BIC, Bayesian information criterion; CFA, confirmatory factor analysis; CFI, comparative fit index; RMSEA, root mean square error of approximation.

Another limitations is that we used only hypothetical gambling tasks; participants were not incentivized to win real money. It has been shown that participants at times behave differently when real money are at stake (for a review, see Kühberger, Schulte-Mecklenbeck, & Perner, 2002). However, the literature favoring real-stakes situations does not generally report very large differences between hypothesized and real-stakes situation. Most importantly, when there are documented differences in effects, they are quantitative in nature (Kühberger et al., 2002). Indeed, our results do not seem to differ from studies that have used incentivized tasks (Frey et al., 2017; Pedroni et al., 2017).

Finally, this study is also limited in regard to presenting a correct model of the data. Although the absolute fit of the CFA was sufficient to conduct model comparisons, the variance explained by the models were low, suggesting that data may be better explained by models that we did not test. However, the low explained variance can also be a result of the notion that there seem to be an overall low reliability and low correlations between measures in the field of psychology; models will not improve their explanatory power as long as we fill them with large amounts of noise. The contamination hypothesis may still be the best model, but it needs to be specified using a selected number of DVs that have high convergent validity and reliability. Although our study involves additional variables and a slightly different focus, it should be noted that the contamination hypothesis model is consistent with the proposed psychometric model of Frey et al. (2017).

4.4 | Conclusions and implications

These results open up for the possibility that many of the RRPs define relatively advanced demands on people's ability to process and understand numbers and abstractions such as "probability," rather than risk preferences per se. This is likely to be especially true when (as we do here) we consider a more general population than the default one of university undergraduates, which may be lower in education, general intelligence, and numeracy. SRPs may provide a more valid measure of risk attitude. This conclusion might appear surprising given the long tradition of "put-your-money-where-your-mouth is" ethos in economic theory, which would appear to assign more validity to actual behavior in real lottery decisions. Our results, along with the recent literature, suggest that RRPs should be applied with great caution until they are better understood.

ORCID

Philip Millroth https://orcid.org/0000-0001-7943-508X

REFERENCES

Agren, T., Millroth, P., Andersson, P., Ridzén, M., & Björkstrand, J. (2019). Detailed analysis of skin conductance responses during a gambling task: Decision, anticipation, and outcomes. *Psychophysiology*, *56*(6), 1–10. https://doi.org/10.1111/psyp.13338

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Icwsm*, *8*, 361–362.

Beauchamp, J. P., Cesarini, D., & Johannesson, M. (2017). The psychometric and empirical properties of measures of risk preferences. *Journal of Risk and Uncertainty*, *54*(3), 203–237.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Becker, A., Deckers, T., Dohmen, T., Falk, A., & Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annual Review of Economics*, *4*(1), 453–478.

Behrens, J. T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods*, *2*(2), 131–160.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review*, *125*(4), 606–615.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258.

Camerer, C. F., Lowenstein, G., & Rabin, M. (Eds.) (2011). *Advances in behavioral economics*. Princeton, NJ: Princeton University Press.

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., & Camerer, C. (2018). *Econographics* (No. w24931). National Bureau of Economic Research.

Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R. (2018). Decision making skill: From intelligence to numeracy and expertise. In *Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Demaree, H. A., DeDonno, M. A., Burns, K. J., & Everhart, D. E. (2008). You bet: How personality differences affect risk-taking preferences. *Personality and Individual Differences*, *44*(7), 1484–1494.

Demaree, H. A., DeDonno, M. A., Burns, K. J., Feldman, P., & Everhart, D. E. (2009). Trait dominance predicts risk-taking. *Personality and Individual Differences*, *47*(5), 419–422.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2018). On the relationship between cognitive ability and risk preference. *Journal of Economic Perspectives*, *32*(2), 115–134.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability?. *American Economic Review*, *100*(3), 1238–1260.

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E. J. (2018). Bayesian inference for Kendall's rank correlation coefficient. *The American Statistician*, *72*(4), 303–308.

Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*, 2319. https://doi.org/10.1038/s41467-019-10301-1

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, *82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, *133*(4), 1645–1692.

Fiedler, K. (2011). Voodoo correlations are everywhere—Not only in neuroscience. *Perspectives on Psychological Science*, *6*(2), 163–171. https://doi.org/10.1177/1745691611400237

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, *3*(10), e1701381. https://doi.org/10.1126/sciadv.1701381

Gillen, B., Snowberg, E., & Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the Caltech Cohort Study. *Journal of Political Economy*, *127*(4), 1826–1863.

Glimcher, P. W., & Fehr, E. (Eds.) (2014). *Neuroeconomics: Decision making and the brain*. New York: Academic Press.

Goekoop, R., Goekoop, J. G., & Scholte, H. S. (2012). The network structure of human personality according to the NEO-PI-R: Matching network community structure to factor structure. *PLoS ONE*, *7*(12), e51558. https://doi.org/10.1371/journal.pone.0051558

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Griffin, A. S., Guillette, L. M., & Healy, S. D. (2015). Cognition and personality: An analysis of an emerging field. *Trends in Ecology & Evolution*, *30*(4), 207–214. https://doi.org/10.1016/j.tree.2015.01.012

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic journal of business research methods*, *6*(1), 53–60.

Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE*, *9*(6), e98679.

Judd, C. M., & Kenny, D. A. (2010). *Data analysis in social psychology: Recent and recurring issues*. Hoboken, NJ, US: John Wiley & Sons, Inc.

Kelder, T., Conklin, B. R., Evelo, C. T., & Pico, A. R. (2010). Finding the right questions: Exploratory pathway analysis to enhance biological discovery in large datasets. *PLoS Biology*, *8*(8), e1000472.

Keren, G., & Wu, G. (Eds.) (2015). *The Wiley Blackwell handbook of judgment and decision making*. United States: John Wiley & Sons.

Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675–709.

Kühberger, A., Schulte-Mecklenbeck, M., & Perner, J. (2002). Framing decisions: Hypothetical and real. *Organizational Behavior and Human Decision Processes*, *89*(2), 1162–1175.

Lai, K., & Green, S. B. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate behavioral research*, *51*(2-3), 220–239.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., … Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84.

Leshem, R., & Glicksohn, J. (2007). The construct of impulsivity revisited. *Personality and Individual Differences*, *43*(4), 681–691.

Levin, I. P., Weller, J. A., Pederson, A., & Harshman, L. (2007). Age-related differences in adaptive decision making: Sensitivity to expected value in risky choice. *Judgment and Decision making*, *2*(4), 225–233.

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Li, Y., Gao, J., Enkavi, A. Z., Zaval, L., Weber, E. U., & Johnson, E. J. (2015). Sound credit scores and financial decisions despite cognitive aging. *Proceedings of the National Academy of Sciences*, *112*(1), 65–69.

Lilleholt, L. (2019). Cognitive ability and risk aversion: A systematic review and meta analysis. *Judgment and Decision making*, *14*(3), 234–279.

Lindskog, M., Kerimi, N., Winman, A., & Juslin, P. (2015). A Swedish validation of the Berlin numeracy test. *Scandinavian Journal of Psychology*, *56*(2), 132–139. https://doi.org/10.1111/sjop.12189

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202.

Lönnqvist, J. E., Verkasalo, M., Walkowitz, G., & Wichardt, P. C. (2015). Measuring individual risk attitudes in the lab: Task or ask? An empirical comparison. *Journal of Economic Behavior & Organization*, *119*, 254–266.

Mamerow, L., Frey, R., & Mata, R. (2016). Risk taking across the life span: A comparison of self-report and behavioral measures of risk taking. *Psychology and Aging*, *31*(7), 711–723. https://doi.org/10.1037/pag0000124

Mata, R., Frey, R., Richter, D., Schupp, J., & Hertwig, R. (2018). Risk preference: A view from psychology. *Journal of Economic Perspectives*, *32*(2), 155–172.

Mata, R., Josef, A. K., Samanez-Larkin, G. R., & Hertwig, R. (2011). Age differences in risky choice: A meta-analysis. *Annals of the New York Academy of Sciences*, *1235*, 18–29. https://doi.org/10.1111/j.1749-6632.2011.06200.x

Mishra, S., Barclay, P., & Sparks, A. (2017). The relative state model: Integrating need-based and ability-based pathways to risk-taking. *Personality and Social Psychology Review*, *21*(2), 176–198. https://doi.org/10.1177/1088868316644094

Mishra, S., & Lalumière, M. L. (2011). Individual differences in risk-propensity: Associations between personality and behavioral measures of risk. *Personality and Individual Differences*, *50*(6), 869–873.

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582.

Nicholson, N., Fenton-O'Creevy, M., Soane, E., & Willman, P. (2002). *Risk propensity and personality*. London: London Business School, Open University Business School and Said Business School Oxford.

Nicholson, N., Soane, E., Fenton-O'Creevy, M., & Willman, P. (2005). Personality and domain-specific risk taking. *Journal of Risk Research*, *8*(2), 157–176.

Pearl, J. (2000). *Causality: Models, reasoning and inference* (Vol. 29). Cambridge: MIT press.

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, *1*(11), 803–809. https://doi.org/10.1038/s41562-017-0219-x

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Poisot, T. (2013). An a posteriori measure of network modularity. *F1000Research*, *2*, 130. https://doi.org/10.12688/f1000research.2-130.v3PMCID:PMC3901540

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociologi-*

cal Methodology, *25*, 111–163. https://doi.org/10.2307/271063
Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(4), 369–379.
Rosenbaum, G. M., & Hartley, C. A. (2018). Developmental perspectives on risky and impulsive choice. *Philosophical Transactions of the Royal Society B*, *374*(1766). https://doi.org/10.1098/rstb.2018.0133
Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.
Sripada, C., Kessler, D., & Jonides, J. (2016). Sifting signal from noise with replication science. *Perspectives on Psychological Science*, *11*(4), 576–578. https://doi.org/10.1177/1745691616652875
Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, *67*(2), 76–90.
JASP Team (2018). JASP (Version 0.9.1) [Computer software].
Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, *34*(1), 23–25.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.
Von Neumann, J., & Morgenstern, O. (1944/2007). *Theory of games and economic behavior (commemorative edition)*. Princeton: Princeton university press.
Walker, D. A. (2003). JMASM9: Converting Kendall's tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods*, *2*(2), 525–530.
Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, *15*(4), 263–290.
Winman, A., Juslin, P., Lindskog, M., Nilsson, H., & Kerimi, N. (2014). The role of ANS acuity and numeracy for the calibration and the coherence of subjective probability judgments. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00851
Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*(4), 493–504.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Millroth P, Juslin P, Winman A, Nilsson H, Lindskog M. Preference or ability: Exploring the relations between risk preference, personality, and cognitive abilities. *J Behav Dec Making*. 2020;1–15. https://doi.org/10.1002/bdm.2171

APPENDIX A: | Network analysis using different correlation matrices

First, we tested if using Pearson's correlation coefficient (*r*) would have an effect on the community structure (using correlations with evidence BF > 10). As illustrated in Figure A1, the same three communities as in the main results were detected (modularity of *Q* = .335): (a) personality measures capturing impulsiveness and individual well-being, (b) stated risk preferences together with a group of personality measures, and (c) cognitive abilities together with revealed risk preferences. Second, we tested if lowering the threshold to including rank-based correlations (i.e., Kendall's tau) with BF > 3 would have an effect on the community structure. As illustrated in Figure A2, the same three communities were again detected (modularity of *Q* = .333).

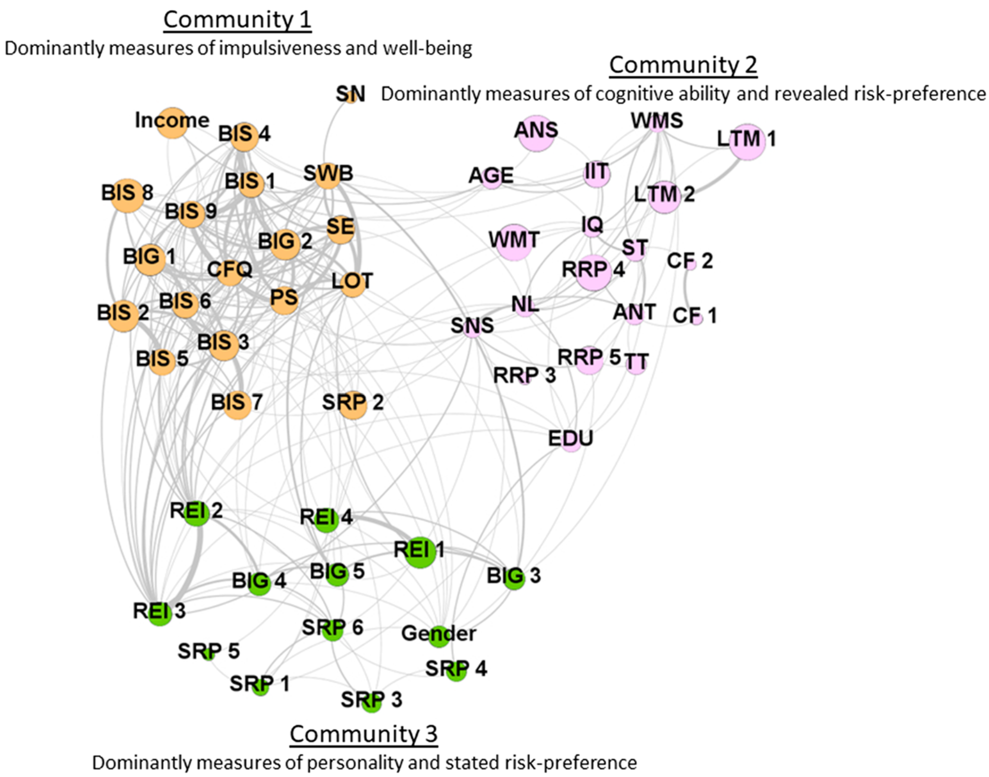

FIGURE A1 Visualization of the relational network where lines constitute links that are constituted by correlations (Pearson's correlations with BF > 10) and nodes (dependent variables [DVs]) are constituted by dots. Size of lines increase dependent on the strength of the correlation. Node size is dependent on how many links are connected to the node: Increased number of connections results in an increased node size. Colors mark the three detected communities (Pink = DVs from measures of revealed risk preference and cognitive ability; Green = dominantly DVs from measures of stated risk preference and personality; Orange = dominantly DVs from measures of impulsivity and well-being). A complete list of abbreviations is provided in Tables 1 and 2. DVs not evident in the figure did not correlate with other DVs [Colour figure can be viewed at wileyonlinelibrary.com]

APPENDIX B: | List of dependent variables included in the confirmatory factor analysis

Table B1 provide the list of dependent variables that were included in the confirmatory factor analysis testing of the three competing hypotheses. Note that the following dependent variables were derived from aggregating the submeasures to an overall mean score for each participant: RPS, LTM, CF, REI 5, REI 6, and BIS.

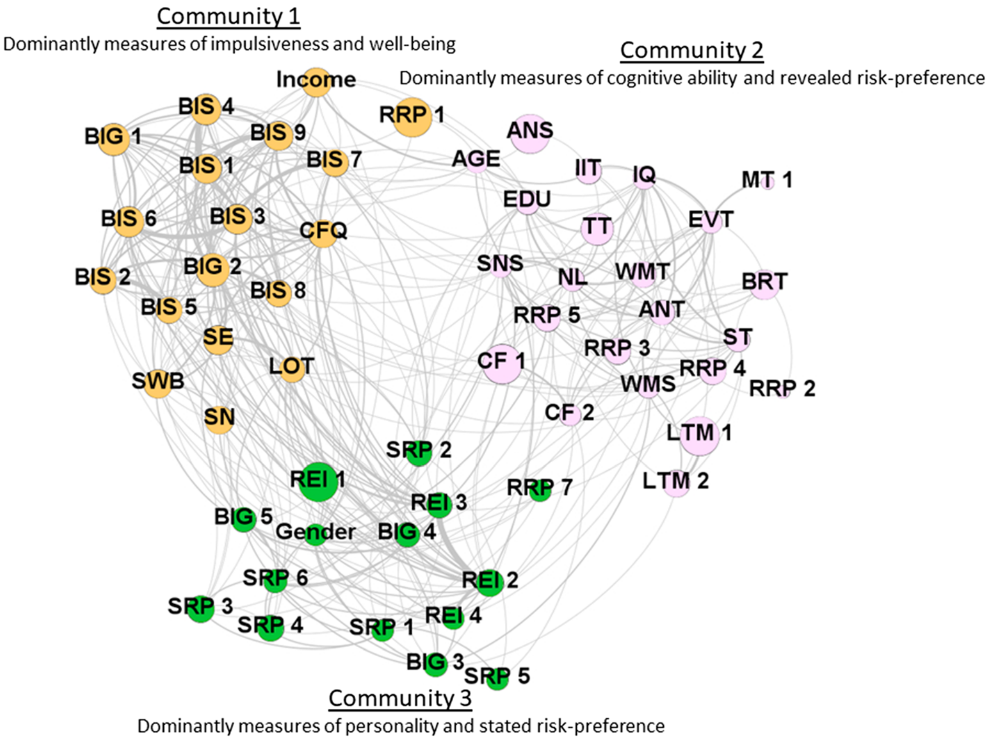

FIGURE A2 Visualization of the relational network where lines constitute links that are constituted by correlations (Tau correlations with BF > 3) and nodes (dependent variables [DVs]) are constituted by dots. Size of lines increases dependent on the strength of the correlation. Node size is dependent on how many links are connected to the node: Increased number of connections results in an increased node size. Colors mark the three detected communities (Pink = DVs from measures of revealed risk preference and cognitive ability; Green = dominantly DVs from measures of stated risk preference and personality; Orange = dominantly DVs from measures of impulsivity and well-being). A complete list of abbreviations is provided in Tables 1 and 2. DVs not evident in the figure did not correlate with other DVs [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE B1 List of DVs included in the confirmatory factor analysis (CFA), detailing the type of test (behavioral or self-report), name of the test, aim of the test, and how the measure is abbreviated in the figures

Types of test	Test name	Aimed to measure	Abbreviation
RRP	Cups Task	Proportion of risky choices	RRP 1
RRP	Discounting	Intertemporal risk attitude	RRP 2
RRP	Iowa Gambling Task	Adaptive risk-taking	RRP 3
RRP	Loss Aversion Task	Attitudes for negative risky prospects	RRP 4
RRP	WTP Task	Attitude to risky prospects	RRP 5
RRP	Wakker Test	Utility function for risky prospect	RRP 6
RRP	Wakker Test	Probability function for risky prospects	RRP 7
SRP	Risk Propensity Scale	Overall risk attitude	SRP-M
Self-report	SNS	Subjective numeracy	SNS
Behavioral	ANS Test	Acuity of mental number line	ANS
Behavioral	ANT	Objective numeracy	ANT
Behavioral	Base Rate Test	Probabilistic reasoning	BRT
Behavioral	Inspection-Time Test	Mental and perceptual speed	IIT
Behavioral	Lipkus	Numeracy	NL

(Continues)

TABLE B1 (Continued)

Types of test	Test name	Aimed to measure	Abbreviation
Behavioral	LTM Test	Overall memory	LTM
Behavioral	Mouselab	Information search	IS
Behavioral	Raven's Matrices	IQ	IQ
Behavioral	Syllogism Test	Reasoning confidence	ST
Behavioral	WM Test	Working memory capacity	WMT
Behavioral	Expected Value Task	Knowledge about EV	EVT
Behavioral	Transitivity Test	Transitive ability	TT
Behavioral	Wechsler Memory Scale	Verbal memory	WMS
Behavioral	Conjunction Fallacy Test	Probabilistic operations	CF
Self-report	Life Orientation Test	Pessimistic and optimistic attitudes	LOT
Self-report	Self-Esteem Test	Self-esteem	SE
Self-report	SWB Test	Subjective well-being	SWB
Self-report	SC Test	Social network	SN
Self-report	PS Test	Perceived stress	PS
Self-report	Big-5 Test	Conscientiousness	BIG 1
Self-report	Big-5 Test	Stability	BIG 2
Self-report	Big-5 Test	Agreeableness	BIG 3
Self-report	Big-5 Test	Openness	BIG 4
Self-report	Big-5 Test	Extraversion	BIG 5
Self-report	Rational Experiential Inventory	Experiential overall	REI-E
Self-report	Rational Experiential Inventory	Rational overall	REI-R
Self-report	Baratt Impulsiveness Scale	Overall impulsiveness	BIS
Self-report	Cognitive Failure Questionnaire	Perceived cognitive ability	CFQ