

Modeling Risk-Taking from Different Measurement Instruments¹

Anna Beatriz Carnielli Howat-Rodrigues², Rosana Suemi Tokumar³, Patrícia Izar⁴

²*Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brazil*

³*Universidade Federal do Espírito Santo, Vitória-ES, Brazil*

⁴*Universidade de São Paulo, São Paulo-SP, Brazil*

Abstract: Risk-taking researches have presented different forms of construction measurements. First, we aimed at evaluating the evidence of validity of the instruments/methods based on external criteria by contrasting the groups on their gender under the three different risk-taking measures: a domain-specific scale and two decision-making tasks in risky situations (Driving a Car game and card games). After that, we aimed at constructing a risk-taking model from the analysis of the relationship between the instruments/methods. The 211 participants (121 women; $M = 21.60$ years old; $SD = 2.19$; and 91 men; $M = 21.46$ years old; $SD = 2.0$) answered the research in a single experimental session with an average duration of 30 min. Individually, the instruments presented criterion validity for contrasting groups. There was no correlation between the scores obtained for the different instruments in the risk-taking model. In conclusion, the data was designed in a three-factor structure indicating the multidimensionality of the risk-taking construct.

Keywords: risk, decision making, risky behavior, measurement, quantitative research

Modelando a Propensão ao Risco a partir de Diferentes Instrumentos de Medida

Resumo: Este estudo teve por objetivo avaliar evidências de validade dos instrumentos/métodos com base em critérios externos, contrastando os grupos pelo sexo nas três diferentes medidas de propensão ao risco: Escala de Propensão ao Risco Específico (EPRE-DE), jogo de dirigir um carro e jogo de cartas. Objetivou-se também construir um modelo de propensão ao risco a partir da análise das relações entre os três instrumentos/métodos. Os 211 participantes (121 mulheres; $M = 21.60$ anos; $DP = 2.19$; e 91 homens; $M = 21.46$ anos; $DP = 2.0$) responderam à pesquisa em uma única sessão experimental com média de duração de 30 minutos. Individualmente, os instrumentos apresentaram validade de critério para grupos contrastantes. Não houve correlações significativas entre os diferentes instrumentos no modelo de propensão ao risco. Concluímos que os dados foram modelados em uma estrutura de três fatores indicando a multidimensionalidade do construto propensão ao risco.

Palavras-chave: risco, tomada de decisão, comportamento de risco, medidas, pesquisa quantitativa

Modelando la Propensión al Riesgo a Partir de Diferentes Medidas

Resumen: La investigación de propensión al riesgo presenta distintas formas de medición del constructo. Primero, se objetivó evaluar evidencias de validez de los instrumentos/métodos sobre la base de criterios externos, contrastando los grupos por el sexo en las tres diferentes medidas de propensión al riesgo: escala de propensión al riesgo, y dos tareas de toma de decisiones. Luego, se objetivó construir un modelo de toma de riesgos a partir del análisis de la relación entre los instrumentos/métodos. Los 211 participantes (121 mujeres; $M = 21.60$ años, $DE = 2.19$; y 91 hombres; $M = 21.46$ años, $DE = 2.0$) respondieron la investigación en una sola sesión experimental con una duración promedio de 30 min. Los instrumentos presentados validez por el grupos de contraste. No se encontraron correlaciones significativas entre los diferentes instrumentos. Los datos se modelan en una estructura de tres factores indicando la multidimensionalidad de lo constructo propensión al riesgo.

Palabras clave: riesgo, toma de decisión, conducta de riesgo, medidas, investigación cuantitativa

¹Paper deriving from the doctoral dissertation with the collaboration of the second author and under supervision of the third author and defended in 2014 in the Postgraduate Program in Psychology at the Institute of Psychology at the University of São Paulo. Support: São Paulo Research Foundation (Fapesp. Grant n° 2011/50645-2).

Correspondence address: Patrícia Izar. Universidade de São Paulo. Instituto de Psicologia. Avenida Professor Mello de Moraes, 1721 - Butantã, São Paulo-SP, Brazil. CEP 05.508-030. E-mail: patizar@gmail.com

Risk-taking refers to the individual's inclination to choose risky options instead of safer ones. The risky option is usually defined by its multiple possible outcomes while the safe option is defined by a single predicted outcome (Helfinstein et al., 2014). Classically, cognitive theorists (e.g. Expected Utility Theory) explain that decisions between safe and risky options are made considering the existence of the best and right choice. Therefore, decisions would

be logically coherent and consistent. As a criticism to this excessive rationality, Prospect Theory researchers propose that decision-making depends on information prospects that involves: how the problem is presented; individual perception of the environmental situation; individual perception of the possibilities for action and outcomes; and the possibility of re-evaluation of the action against its outcomes (Howat-Rodrigues, Andrade, & Tokumaru, 2013; Linde & Vis, 2016).

Currently, researchers interested in risk-taking have added several factors to the analyses of the phenomenon such as the domains in which decision-making occurs (e.g. social, financial), the psychological processes involved (e.g. affective, deliberative), the perception of gains and losses, and the influence of individual features such as personality traits, individual perception of risk involved in a task and gender (Howat-Rodrigues et al., 2013). This scenario leads to the employment of different techniques for measuring risk-taking and difficulty in standardizing these instruments and establishing convergent validity among them (Boyer, 2006).

The literature presents a wide variety of methods used to measure risk-taking. Systematic reviews (Boyer, 2006; Harrison, Young, Butow, Salkeld, & Solomon, 2005) showed that previous research employed both risk-taking scales and tasks. Harrison et al. (2005) listed a total of 12 instruments to measure risk-taking, 11 of them were scales. The authors identified two approaches to measure risk-taking: through associated personality traits and across multiple specific domains.

Scales that measure associated personality traits are consistent with the assumption that risk-taking is a stable personality trait related to general traits and motivational tendencies (Blais & Weber, 2006). Examples of scales developed within this assumption are the Arnett's Inventory of Sensation Seeking – AISS (Arnett, 1994) with two subscales - Intensity (Cronbach's alpha = .64) and Novelty (Cronbach's alpha = .50) and the Sensation Seeking Scale – SSS (Zuckerman, Eysenck, & Eysenck, 1978) with three subscales – Thrill and Adventure, Disinhibition, Boredom Susceptibility (reliabilities ranged from .56 to .82). Risk-taking is considered to be positively correlated with the search for sensation and/or danger and impulsivity and negatively correlated with self-control.

The measurement of risk-taking across domains is consistent with the assumption that individual perception of the context strongly influences risk-taking (Blais & Weber, 2006). Examples of scales within this assumption are the Specific Risk-Taking Scale (Howat-Rodrigues et al., 2013) with two subscales: Cohesion and Competition (reliabilities established for the subscales ranged from .66 to .83 for different subsamples) and the Domain-specific Risk-Taking Scale – DOSPERT (Weber, Blais, & Betz, 2002) with five subscales: Finance, Health/Safety, Recreational, Ethical and Social (reliabilities established for the subscales ranged from .67 to .89). These scales measure the individual risk perception or/and attitude in a Likert scale (how likely he/she would be to perform the behaviour).

Previous studies showed correlation amongst risk scales evaluated through personality traits (e.g. correlation between AISS and SSS was .41, $p < .001$) (Arnett, 1994); and between these scales and the Balloon Analogue Risk Task – BART, which is a computerized game used to measure risk-taking, BART and Eysenck Impulsivity subscale, $r = .24$; $p < .05$ (Harrison et al., 2005); Bart and SSS, $r = .35$; $p < .01$ (Boyer, 2006). Both measurements of personality traits and measurements of specific domains present valid instruments for the study of risk-taking. In summary, sensation seeking and impulsivity are personality traits used as predictors of higher risk-taking. While, the DOSPERT scale measures risk perception and allows for the statistical inference of the probability of a person having a risky attitude (Harrison et al., 2005).

Scalar instruments require the participant to report what he/she would do in a hypothetical situation. The consideration of the social rules and of what is socially acceptable is implicit in the response to such instruments (Lönnqvist, Verkasalo, Walkowitz, & Wichardt, 2015). Otherwise, instruments based on tasks require decision-making for the proposed situations (Helfinstein et al., 2014). Decisions lead the participant to a result (grades or monetary rewards) that allows him/her to win a prize. Lönnqvist et al. (2015) suggested that risk-taking displayed on specific tasks, such as games linked to a given context, should be different from propensity assessed by cognitive tasks such as scales. The authors argue that tasks are more likely to engage the attention and motivation of the participant. In general, the measurement of risk-taking in a more contextual manner involves game performance. The attribute that both tasks have in common is decision-making when facing unpredictable contexts (Mather et al., 2012). However, there are differences according to the characteristics involved in decision-making.

We identified three major groups of risk-taking tasks: (a) go/no go games; (b) gambling games (gambling tasks); and (c) delay of gratification tasks. Within the group of go/no go games we added BART and the “Chicken” - Driving a Car game (Mather et al., 2012). Overall, these games consist of performing a task to earn more points (e.g., to move a cart or fill up a balloon). However, the participant loses points if he/she performs the task in the presence of an inhibitory signal that appears at unpredictable times during the procedure. Go/no go games included elements of impulsivity control and discrimination of stimuli.

Delay of gratification tasks are also related to impulsivity control. These tasks are used to measure the preference for earning a smaller, more immediate reward than winning a larger reward later on. The methods are applied through scales or procedures with quantitative, monetary choices or flavoursome rewards such as marshmallow and chocolate (e.g. Stanford marshmallow experiment). In a longitudinal study, Casey et al. (2011) found that individuals who were less able to delay gratification in preschool (marshmallows task) consistently showed low self-control abilities (suppress actions in go/no-go tasks) in twenties and thirties.

Gambling games are based on the presentation of 2 or more choices (e.g. decks of cards, boxes or doors). Each choice has a possibility of gain and/or loss different from one another within a predetermined scheme. Typically, the participants have the opportunity to take samples of cards (or open doors or boxes) from all sets of choices in the order they want, thus assessing, throughout the games, what is the probability of gain and/or loss of each set. Examples of this kind of task are the Iowa Gambling Task - IGT, and the card game developed by Weber, Shafir, and Blais (2004), detailed in method.

Boyer's systematic review (2006) identified 111 studies that investigated the development of risk-taking in four perspectives: cognitive ($n = 33$), emotional ($n = 20$), psychobiological ($n = 25$) and social ($n = 33$). The studies involved different methods such as observational, experimental, interview, and physiological assessment. Here, we highlight the results of seven studies that mixed experimental and interview methods. Five of them used personality scales (e.g. Eysenck Impulsiveness Scale-V, Personality Inventory) and go/no-go (e.g. BART) or gambling tasks (e.g. IGT). Overall, results were inconsistent and few correlations were found, especially among gambling tasks and scales, and among gambling tasks and go/no-go tasks. Rupp et al. (2016) investigated the impact of neurocognitive impulsivity in forty-three alcohol-dependent patients. They used two go/no go tasks (one experimental version available in the psytest.net website and the STOP-IT task), one delay of gratification task (Delay Discounting Test - DDT), and one gambling task (IGT). There was statistically significant correlation only between DDT and IGT ($r = -.38$; $p = .013$). One hypothesis to these results is that the different tasks (e.g. response to questionnaires, response to go/no-go games, and response to gambling task) can be related to the activation of different parts of the brain. The functional Magnetic Resonance Imaging (fMRI) study by Monchi, Petrides, Petre, Worsley, and Dagher (2001) found that the orbital prefrontal cortex and striatal circuits were involved during different stages of the gambling task (Wisconsin Card Sort Task - WCST) performance. Casey et al. (2011) acknowledged the importance of the prefrontal cortex in go/no-go tasks, which showed different activation between no-go and go trials.

Another hypothesis, which does not exclude the one cited above, refers to the multidimensional nature of risk-taking. According to this hypothesis each risk-taking measure could be related to a different risk dimension and individual risk-taking would not have necessarily consistent rates through all contexts. This hypothesis indicates the relevance of the multifactorial model of risk-taking (Howat-Rodrigues et al., 2013; Lönnqvist et al., 2015). However, authors have been finding difficulty in designating and standardizing risk-taking dimensions. For example, Rupp et al. (2016) named all tasks instruments (two go/no go tasks, one delay of gratification task and one gambling task) as impulsivity. In the DOSPERT, Weber et al. (2002) designated two different dimensions: risky attitude and risk perception, on the other

hand, Harrison et al. (2005) is more specific discussing risk propensity, risky attitude, risk perception, decision-making, and others.

Given the complexity of the phenomenon and the inconsistent results found in the presented studies, further investigation of the relationship among different measures of risk-taking can be helpful in the description of the dimensions involved. Some relationships among different instruments have also been poorly explored such as those among domain-specific risk-taking scales and others risk-taking task instruments. This study presented two complementary objectives. First, we evaluated evidence of validity of instruments/methods based on external criteria by contrasting the groups on their gender under three different measures of risk-taking (a domain-specific scale, a go/no-go task and a gambling task). Second, we aimed at constructing a risk-taking model through the analysis of the relationship among the three instruments/methods.

Regarding the first objective, Chan (2014) points out that it is important to verify aspects of validity of the instruments in the studied sample because this relates to the quality of inferences and statistic decisions from the scores of an instrument. We choose gender as variable criterion since previous research recurrently indicates gender patterns of risk-taking. We work from the hypothesis that risk-taking instrument should differentiate risk scores between men and women. Men will show greater risk-taking scores than women in all instruments/methods (Cobey, Stulp, Laan, Buunk, & Pollet, 2013).

In evolutionary terms, gender differences to risk-taking can be explained from Trivers' Theory of Parental Investment. The author proposes that females' reproductive success depends on parental investment in order to ensure that their children reach reproductive age and all of their effort has not been lost since females are responsible for the whole process of gestation and lactation. Because maternal investment is long-term, women tend to be more risk-averse, given high-energy costs of taking risks, whereas males' reproductive success is limited by the number of partners that they can maintain sexual intercourse, so men invest less in parental care than women and can increase their reproductive potential by increasing numbers of potential partners. These biologically limiting differences for reproductive success favoured men more concerned with intrasexual competition as a way of increasing the number of partners, thus being more risk-taking in general (Cobey et al., 2013).

Furthermore, we test the theoretical model presented in Figure 1. Our initial model derives from the literature hypothesis that showed all three of these instruments being used as measure risk-taking and pointed out some correlations between different types of risk-taking measures (Boyer, 2006; Rupp et al., 2016). As for the direction of these relations, positive significant correlations between different measuring methods must indicate that the same context/dimension of risk-taking is being assessed, whereas no significant correlations must indicate the evaluation of different contexts/dimensions.

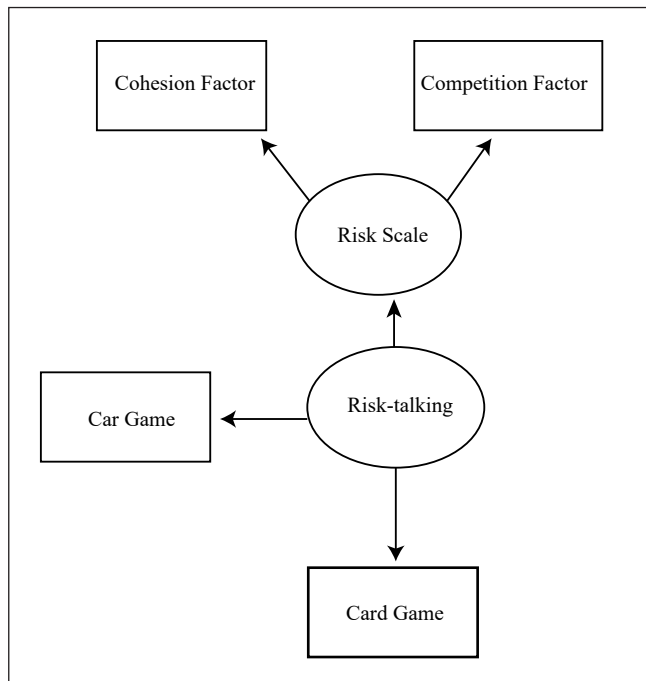


Figure 1. Theoretical model of the relationships of three different instruments used in risk-taking research to measure this construct.

Method

Participants

A total of 211 people participated in the study. They were residents of two states in south-eastern Brazil. A total of 121 participants were women with an average age of 21.60 years old ($SD = 2.19$), and 91 were men with an average age of 21.46 ($SD = 2.0$). Participants were recruited in public universities in the capital of each state throughout advertisements affixed in various locations. Socioeconomic status was measured according to Brazil Economic Classification Criteria: 36.5% (77) were class A, 48.3% (102) were class B, and 15.2% (32) were class C.

Instruments

Risk-taking was measured using 3 different instruments: *Specific Risk-Taking Scale – Evolutionary Domains* (Escala de Propensão ao Risco Específico – Domínios Evolutivos) – EPRE-DE (Howat-Rodrigues et al., 2013). This instrument consists of 14 items self-rated in a 5-point Likert scale (1 = would never do; 5 = would always do) with 2-factor multidimensionality: 6 items distributed in the Cohesion factor (Cronbach's alpha = .75) and 8 in the Competition/Fertility factor (Cronbach's alpha = .66).

The Chicken – Driving a Car game (Fattori, Howat-Rodrigues, & Izar, 2017; Mather et al., 2012). This game is a computer game in which the participant earns points by driving a car while the light is green and stopping it when it is

red. The participant drives the car but does not control its speed. The appearance of a yellow light signals the imminent appearance of a red light and the loss of all points if the car continues moving when the red light is displayed. The computer records the number of stops and the amount of time during which the car is in motion between the appearance of the yellow light and the last stop. Each participant plays 15 rounds (Cronbach's alpha = .91). We use the mean movement time in the rounds as a measure of risk-taking.

Card game: This game was developed by the authors based on the game of Weber et al. (2004). For each participant, two sets of 26 cards were presented. One set always had the same card (x), while the other had two different cards (0 and y , $y > x$). The experimenter provided the following information to the participant: "One set always has a fixed card, and the other always has different cards, zeros and another number. Your goal is to score 21 points in every round". After the instructions, the participant was asked to choose only one of the sets to play in each round. Each round began with the experimenter giving the participant a card with a number equivalent to the number of initial points of that round. The participant would, then, take 3 cards from the set he/she had chosen. If the sum of the points were 21, he/she would earn double the points for that round, but if the sum were more than 21 points, then he/she would not score any points. A total of 8 rounds were performed. The downside risk index ($\delta(k)$) was used as a measure of risk-taking. This index measures the return rate of each choice. The higher and more positive the downside risk index is, the higher the return rate of the choice and thus the lower the risk-taking (Gibson, Seiler, & Walden, 2016). The reliability was estimated by frequency of risk-taking or risk-aversion choice in each trial, KR-20 was .54.

Procedure

Data collection. Participants responded individually in an experimental setting with mean session duration of 30 minutes. With the aim of engaging participants in the games the experimenter told them that they were competing for a prize that would be awarded to the participant if his/her total score was greater than that of the previous participants. The potential prizes, various kinds of chocolates and candies were presented and the participant chose a reward before the task started. The prizes were where participants could see for the entire procedure. This procedure sought to offer a possibility of real gain, simulating decision-making in a natural situation.

Data analysis. First, descriptive statistical calculations with all of the variables involved and correlation tests between risk measures were performed to examine correlations between factors/items. A p value (significance) $< .05$ was considered statistically significant. We analysed the validity by contrasting groups in each instrument using different analyses of variance (ANOVAs). We used the participant's gender as the independent variable and the risk-taking scores in each instrument as dependent variables.

Each item of the scale and the response in each round of the two decision-making tasks were used as observed variable resulting in risk-taking latent variable for each instrument/method. In turn, these instrument latent variables were used in a structural equation modelling analysis (SEM) to verify the emergence of the latent variable of risk-taking with confirmatory modelling strategy. According to the suggestions of Byrne (2012), the following indexes were analysed for the confirmatory factor analysis (CFA): (1) χ^2 (chi-square) – model fit index (values lower than 5, which are not significant, are recommended); (2) χ^2/df – fit indicator (values between 2 and 5 are recommended); (3) CFI (Comparative Fit Index) – a comparative indicator of the model fit (values greater than 0.9 are recommended); (4) RMSEA (Root Mean Square Error of Approximation) – indicator index of residuals suitable for confirmatory strategies of large samples (values lower than .08 at the 90% confidence interval are recommended); (5) GFI (Goodness-of-Fit Index) – index that relates to the proportion of variance-covariance in the data explained by the model through estimation (values greater than .90 are recommended).

Analyses were performed using SPSS (Statistical Package for Social Sciences; IBM Corp©) version 16.0 and Amos (Analysis of Moment Structures) version 7.0.

Ethical Considerations

The Ethics Committee on Human Research (Health Science Center of the Federal University of Espírito Santo – nº 198/11) authorized all stages of this research and all participants signed an informed consent form.

Table 1

Fit indexes of the risk-taking models tested

Model	Fit indices							
	χ^2	d.f.	<i>p</i>	χ^2/df	GFI	CFI	RMSEA (90%CI)	
Intermediate 1	Theoretical Model	829.20	556	< .001	1.49	.81	.87	0.05(0.04-0.06)
Intermediate 1.1	Same as the previous with control of the parameters of e24-e25(34.57)	792.09	555	< .001	1.43	.83	.89	0.05(0.04-0.05)
Intermediate 1.2	Same as the previous with control of the parameters of e2-e3(15.79)	762.68	554	< .001	1.38	.83	.90	0.04(0.04-0.05)
Intermediate 2	Except Competition/fertility factor	526.05	321	< .001	1.64	.84	.89	0.06(0.05-0.06)
Intermediate 2.1 – Final	Same as the Intermediate 2 model with control of the parameters of e24-e25(34.57)	488.94	320	< .001	1.53	.85	.91	0.05(0.04-0.06)

Note. e = variance error associated with each observed variable; χ^2 = Chi-square; χ^2/df – fit indicator; GFI = goodness of fit index; CFI = comparative fit index; RMSEA = Root mean square error of approximation.

Results

Validity evidence based on external criteria by contrasting groups

Gender differences were found in all of the used instruments. In the Driving a Car game, men ($M = .70$, $SD = .15$) presented higher mean time of car movement [$F(1, 209) = 6.01$; $p = .02$; Cohen's d ($d = .36$)] than women ($M = .64$, $SD = .18$) ranging from .10 min to 1.01 min. In the Card game, men ($M = 2.91$, $SD = .84$) also showed higher means on the downside risk index [$F(1, 209) = 4.81$; $p = .03$; $d = .31$] than women ($M = 2.62$, $SD = 1.05$) ranging from 0 to 3.85. In the risk-taking scale (EPRE-DE), the differences occurred for the dimension of cohesion [$F(1, 209) = 6.74$; $p = .01$; $d = .36$] with men ($M = 1.57$, $SD = .62$) obtaining higher mean values than women ($M = 1.38$, $SD = .43$). However, there was no significant difference in the dimension of competition/fertility ($M_{women} = 1.95$, $SD = .51$, $M_{men} = 1.90$, $SD = .48$; $d = .10$). In addition, statistically significant differences had low effect size. Both dimensions of the EPRE-DE ranged from 0 to 5. Means of men and women in each measure showed that in the scale the means drifted towards the minimum value whereas while in the games the means tended towards the maximum values.

Relationships among risk measurements

The observed variables – scores in the subscales cohesion and competition/fertility, mean time of car movement, and downside risk index – were loaded into the confirmatory factor analysis (CFA) using the Maximum Likelihood method to assess the existence of the latent variable of risk-taking. Table 1 shows the index of all intermediate models up to the selection of the final model, which showed acceptable indexes.

First, we considered the variables in the theoretical model presented in Figure 1 (Table 1 – intermediate model 1, 1.1, and 1.2). Seeking to improve the adjustment of the model, we analyzed the modification indices (MI) in order to identify suggestions of correlation among the error (residue) parameters of pairs of items or the existence of crossed loadings with indices that showed values above 11 being analyzed. An elevated MI was found in the relationships $e_{24} - e_{25} = 34.57$ and $e_2 - e_3 = 15.79$ (intermediated model 1.1; 1.2). The control parameters did not contribute to the overall fit of the model and we opted for the exclusion of the competition/fertility (intermediate model 2) variable. This decision-making was also influenced by the fact that: (a) this factor did not show validity differences by contrasting group, (b) and previous researches showed best fit of the scalar instrument at one factor (Howat-Rodrigues et al., 2013).

All χ^2 indexes were high and significant (Table 1). According to Byrne (2012), these values indicate a lack of model fit; however, the same authors claim that for the final decision on model fit other indexes should be analysed. If the χ^2 value is significant, it is divided by the degrees of

freedom (χ^2/df). The model is acceptable when it presents values equal to or lower than 5. All proposed models in this study showed values within this range. Moreover, the descriptive and restrictive indexes of model fit, i.e., GFI, CFI, and RMSEA, were also satisfactory, showing the best values in the final model (intermediate 2.1) indicated in Table 1. All observed variables showed significant regression weight. In final model, GFI index showed the best fit, although still below the reference value ($> .90$). Byrne (2012) recommend analysing all the indices together for the decision on the statistical adequacy of the model, and in the general, our indexes were satisfactory, so we analysed this as a final decision since GFI index below the reference value.

The final model (Figure 2) indicated the existence of an acceptable model of risk-taking. The data was modelled in a three-factor structure. The cohesion variables, downside risk index, and mean time of car movement did not correlate (cohesion-downside risk index, $r = .13$; $p = .20$; cohesion-mean time of car movement, $r = .06$; $p = .50$; mean time of car movement-downside risk index, $r = .17$; $p = .12$). Figure 2 shows the standardized estimates.

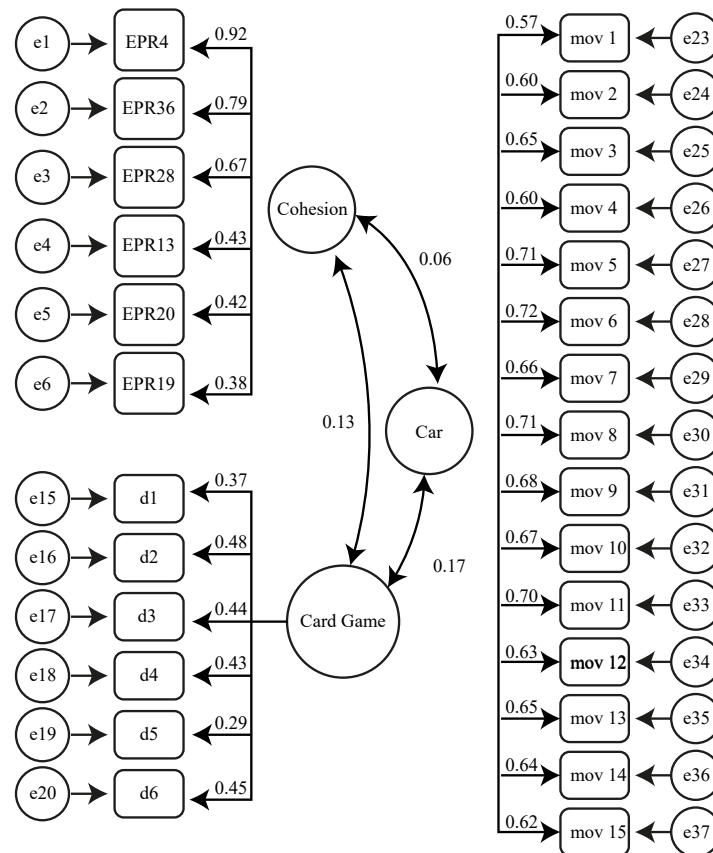


Figure 2. Confirmatory Factor analysis model of the dimensions of risk-taking showing standardized estimates. Regression coefficients are indicated near the 1-way arrows, and multiple correlations to the square are indicated near the variables. EPR4, EPR13, EPR19, EPR20, EPR28, and EPR36 correspond to the items of the cohesion factor of the Specific Risk-Taking Scale - Evolutionary Domains; d1-d6 correspond to the downside index shown in moves 1-6 in the card game – moves 3 and 4 were excluded due to their high return rates and thus low risk; mov 1-15 correspond to the mean movement time in rounds 1-15 in the Driving a Car game.

Discussion

One of our initial hypotheses was that a valid risk taking measure should differentiate scores from men and women. Men should show greater risk-taking scores than women as gender is seen as an important factor for variation in risk-taking (Cobey et al., 2013). The results presented here indicated that all used risk measures were valid in terms of the contrasting groups, confirming results from previous research studies. Regardless of the effects, the intensity was low; Nevertheless, it might still be relevant (Espirito-Santo & Daniel, 2015) due to the difficulty in standardizing risk-taking measures (Boyer, 2006) and a greater human inclination to risk aversion (Linde & Vis, 2016; Weber et al., 2004).

The lack of difference between risk-taking scores of men and women in one of the factors (competition/fertility) of the domain-specific scale and the lack of fit of that factor to the final model can be attributed to the statistical instrument's fragility as observed by the authors during the validation of the instrument (Howat-Rodrigues et al., 2013). Recurrently, Brazilian versions of the EPRE result into one factor with good statistical fit that aggregate the items related to the use and abuse of substances (Howat-Rodrigues et al., 2013). In the version presented here, these items were grouped in the cohesion factor. In all versions of the EPRE, other factors presented weak statistical fit, which was repeated in the results reported here for the competition/fertility factor.

The data presented a possible risk-taking model, modelled in a three-factor structure, and each measurement did not present correlation amongst them, behaving as a specific factor. This result agrees with others (Casey et al., 2011; Lönnqvist et al., 2015; Monchi et al., 2001; Rupp et al., 2016) that also found no correlation among different risk-taking instruments, and neuroimaging studies (Casey et al., 2011; Monchi et al., 2001) indicated the activation of different brain areas in face of different tasks related to risk-taking. These convergences of results can indicate the multidimensional nature of risk-taking. This inference is in accordance with our initial proposition that significant correlations between different risk-taking measures should indicate that the same context/dimension of risk-taking was being assessed, whereas no significant correlations should indicate the evaluation of different contexts/dimensions (Howat-Rodrigues et al., 2013; Lönnqvist et al., 2015). It is also in accordance with psychological models that disagree with the view of risk-taking as a general trait of the individual and emphasize the relationship between risk-taking and loss/gains in particular contexts (Howat-Rodrigues et al., 2013; Weber et al., 2002).

Analysing individually the instruments, we found the following: the scale consistently evaluated the cohesion domain because the competition/fertility domain was eliminated from the analysis. Although only one dimension was consistently evaluated, the task required from the participants to answer the scale to report their perception of risk for each behaviour presented in the items. Losses and gains were not explicitly declared and unpredictability was not manipulated. Based on the analysis of Harrison et al. (2005), we suggest that the

EPRE-DE used here is measuring a dimension of risk taking that can be described as the participants' risky attitude.

By contrast, in the two tasks, participants experienced situations of unpredictability (situations with no explicit probabilities) and risk (situations in which there were gain variance). However, both games were related to different dimensions of risk-taking. In the Driving a Car game, the participant had no information about the moment in which the negative consequence would appear (the red light and loss of points), and the choice with less risk involved keeping the points already accumulated in the game. Although this game has features that could be fitted into risky attitudes, the data showed that it is part of a different dimension of the one proposed to the scalar instrument. This dimension could be described as risk propensity (Harrison et al., 2005).

In the card game, the participant had to choose between a set that always offered the same card (high predictability) and another set that offered at least two types of cards, 0 and another value (low predictability). This procedure involved less unpredictability than the Driving a Car game because the participant could opt for the most predictable option, better resembling investment risks. We could describe this dimension of risk as decision-making under risk (Harrison et al., 2005).

In both tasks, the majority of the participants chose to take a risk. However, studies show that humans are averse to risk, as are other animals when they are not in danger (i.e., starving) (Weber et al., 2004). Linde and Vis (2016) explain that, according to the Prospect Theory, risk aversion occurs more predictably in gain situations. For example, when people prefer certain gains (\$300) to higher risk options (80% probability of winning \$400), even though the highest risk option has a higher expected value ($80\% \times 400 = \$320$). By contrast, in situations that involve loss, as is the case in the two proposed games, there seems to be an adjustment of strategies. One of the most accepted explanations for this adjustment comes from the Prospect Theory. Intuitively, people judge the possibilities of results from how much is earned and how much is lost in each option. However, because they are more sensitive to losses and disadvantages (averse to losses) than to gains and advantages, they tend to be risk averse in gain situations to avoid possible losses. But, in loss situations people tend to be prone to risk recovering losses.

An important variable that must be considered regarding games is the influence of competition. In the instrument proposed by Weber et al. (2004), for example, the participant earned in cash the value of one of the choices he/she had made. In this case, there was a safe return regardless of the final score earned. However, the resolution of the National Health Council No. 466 (Ministério da Saúde, 2012), which regulates research involving human beings in Brazil, does not allow gains from participation in research studies. Thus, we adapted the application of the instrument by offering the possibility of gain if the participant reached the best score among all the people who had previously played the game, introducing a competition effect, which we did not preview. To attempt to minimize the effect of competition,

the participant was informed that he/she needed to earn more points than all the people who had previously played the games. Nonetheless, the participant was not informed about the maximum score of the ranking, meaning that he/she had incomplete information for making a decision.

At least three aspects of risk-taking were evaluated in this study: risky attitude, risk propensity and risk in decision-making. Although all of them are theoretical linked to the risk-taking construct, we demonstrate that the measures generated different evaluations of risk-taking. The results presented here support the studies of risk-taking as a multidimensional approach (Howat-Rodrigues et al., 2013; Lönnqvist et al., 2015). We indicated that further studies should include personality scales. Personality traits can be used as predictors of higher risk-taking (Harrison et al., 2005) and they could be used to help to standardize risk-taking dimensions and contribute to validity of the measurements.

Participants were all university students, who were submitted to the same conditions and answered to all the three instruments that were evaluated. However, the homogeneity of the sample and the small number of instruments analysed (three) should be considered as limitations of this study. As well as non-inclusion of other variables such as personality traits, age variation of the participants, the impact of physiological, cognitive, motor or emotional variables that have been pointed, in the literature, as important factors to variation in risk-taking (Howat-Rodrigues et al., 2013; Linde & Vis, 2016). In addition, the expansion of the number of risk-taking instruments/method jointly analyzed in future research would enable the refinement of the possible links between different measures of risk-taking, and the standardization of labels for each risk-taking dimension. Surveys that contemplate as many variables as possible can contribute more consistently to the test of competing statistical models, refining the theory in the area of risk-taking.

In this study, we used different instruments to assess different aspects of risk-taking in the same participants. The instruments did not present correlation among them, but they were considered valid in terms of external criteria by contrasting gender groups. The data presented a possible model with a three-factor structure. The development of different instruments, including scalar and simulation measures, which allow the evaluation of different aspects of risk-taking, appears to be a promising research path in the study of the variation of risk-taking in different contexts.

References

- Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, 16(2), 289-296. doi:10.1016/0191-8869(94)90165-1
- Blais, A. R., & Weber, E. U. (2006). *Testing invariance in risk taking: A comparison between Anglophone and Francophone Groups*. Montréal, Canada: CIRANO. Retrieved from <http://www.cirano.qc.ca/pdf/publication/2006s-25.pdf>
- Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review*, 26(3), 291-345. doi:10.1016/j.dr.2006.05.002
- Byrne, B. M. (2012). *Structural equation modeling with Mplus*. New York, NY: Routledge.
- Casey, B. J., Somerville, L. H., Gotlib, I. H., Ayduk, O., Franklin, N. T., Askren, M. K., ... Shoda, Y. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences of the United States of America*, 108(36), 14998-15003. doi:10.1073/pnas.1108561108
- Chan, E. K. H. (2014). Standards and guidelines for validation practices: Development and evaluation of measurement instruments. In B. D. Zumbo & E. K. H. Chan (Eds.), *Validity and validation in social, behavioral, and health sciences* (pp. 9-24). New York, NY: Springer.
- Cobey, K. D., Stulp, G., Laan, F., Buunk, A. P., & Pollet, T. V. (2013). Sex differences in risk taking behavior among Dutch cyclists. *Evolutionary Psychology*, 11(2), 350-364. doi:10.1177/147470491301100206
- Espirito-Santo, H., & Daniel, F. (2015). Calcular e apresentar tamanhos do efeito em trabalhos científicos (1): As limitações do $p < 0,05$ na análise de diferenças de médias de dois grupos [Calculating and reporting effect sizes on scientific papers (1): $p < 0.05$ limitations in the analysis of mean differences of two groups]. *Revista Portuguesa de Investigação Comportamental e Social*, 1(1), 3-16. doi:10.7342/ismt.rpics.2015.1.1.14
- Fattori, K. A., Howat-Rodrigues, A. B. C., & Izar, P. (2017). Risk-taking game: Construction, validity evidence and gender differences. *Psicologia: Teoria e Prática*, 19(1), 137-151. doi:10.5935/1980-6906/psicologia.v19n1p137-151
- Gibson, S., Seiler, M. J., & Walden, E. (2016). Using experimental and neurological data to gain a deeper understanding of realization utility theory. *Real Estate Economics*, 44(2). doi:10.1111/1540-6229.12164
- Harrison, J. D., Young, J. M., Butow, P., Salkeld, G., & Solomon, M. J. (2005). Is it worth the risk? A systematic review of instruments that measure risk propensity for use in the health setting. *Social Science & Medicine*, 60(6), 1385-1396. doi:10.1016/j.socscimed.2004.07.006
- Helfinstein, S. M., Schonberg, T., Congdon, E., Karlsgodt, K. H., Mumford, J. A., Sabb, F. W., ... Poldrack, R. A. (2014). Predicting risky choices from brain activity patterns. *Proceedings of the National Academy of Sciences of the United States of America*, 111(7), 2470-2475. doi:10.1073/pnas.1321728111

- Howat-Rodrigues, A. B. C., Andrade, A. L., & Tokumaru, R. S. (2013). Escala de Propensão ao Risco Específico – Domínios Evolutivos (EPRE-DE): Adaptação e estrutura fatorial. *Avaliação Psicológica*, 12(3), 281-290. Retrieved from http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712013000300002
- Linde, J., & Vis, B. (2016). Do politicians take risks like the rest of us? An experimental test of prospect theory under MPs. *Political Psychology*, 38(1), 101-117. doi:10.1111/pops.12335
- 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. doi:10.1016/j.jebo.2015.08.003
- Mather, M., Mazar, N., Gorlick, M. A., Lighthall, N. R., Burgeno, J., Schoeke, A., & Ariely, D. (2012). Risk preferences and aging: The “certainty effect” in older adults’ decision making. *Psychology and Aging*, 27(4), 801-816. doi:10.1037/a0030174
- Ministério da Saúde. Conselho Nacional de Saúde. (2012). *Resolução No. 466, de 12 de dezembro de 2012* [Resolution No. 466, 2012 December 12]. Aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Retrieved from http://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html
- Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 21(19), 7733-7741. doi:10.1523/JNEUROSCI.21-19-07733.2001
- Rupp, C. I., Beck, J. K., Heinz, A., Kemmler, G., Manz, S., Tempel, K., & Fleischhaker, W. W. (2016). Impulsivity and alcohol dependence treatment completion: Is there a neurocognitive risk factor at treatment entry? *Alcoholism: Clinical and Experimental Research*, 40(1), 152-160. doi:10.1111/acer.12924
- 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. doi:10.1002/bdm.414
- Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk-sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, 111(2), 430-445. doi: 10.1037/0033-295X.111.2.430
- Zuckerman, M., Eysenck, S. B. G., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age and sex comparisons. *Journal of Consulting and Clinical Psychology*, 46(1), 139-149. doi:10.1037//0022-006X.46.1.139
- Anna Beatriz Carnielli Howat-Rodrigues* is a Ph.D. candidate of the Instituto de Psiquiatria at Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brazil.
- Rosana Suemi Tokumaru* is a Professor of the Universidade Federal do Espírito Santo, Vitória-ES, Brazil.
- Patrícia Izar* is a Professor of the Universidade de São Paulo, São Paulo-SP, Brazil.
- Authors’ Contribution:*
All authors made substantial contributions to the conception and design of this study, to data analysis and interpretation, and to the manuscript revision and approval of the final version. All the authors assume public responsibility for content of the manuscript.
- Received:* Jun. 23, 2016
1st Revision: Nov. 16, 2016
2nd Revision: Jun. 29, 2017
3rd Revision: May. 19, 2018
Approved: Jun. 07, 2018
- How to cite this article:*
Howat-Rodrigues, A. B. C., Tokumaru, R. S. & Izar, P. (2018). Modeling risk-taking from different measurement instruments. *Paidéia (Ribeirão Preto)*, 28, e2828. doi: <http://dx.doi.org/10.1590/1982-4327e2828>