

## Pilot Testing and Preliminary Psychometric Validation of Situational Judgment Test of Resistance to Sunk Cost

Rahmat Hidayat<sup>\*1</sup>, Muhammad Dwirifqi Kharisma Putra<sup>1</sup>

<sup>1</sup>*Faculty of Psychology, Universitas Gadjah Mada, Indonesia*

Submission 24 March 2022 Accepted 18 October 2021 Published 23 December 2022

**Abstract.** The sunk-cost fallacy is an anomaly in decision-making that has been proven in various experimental studies. However, individual differences in the tendency to fall into sunk-cost fallacy have not been sufficiently reported. This preliminary study contains a psychometric evaluation report of the Resistance to Sunk Costs (RtSC) measurement instrument, a component of the Adult Decision Making Competence instrument (A-DMC) (de Bruin et al., 2007) which is modified into Indonesian socio-cultural and economic contexts. The RtSC instrument uses a situational judgment test model, therefore it is adequate to measure the psychological construct of sunk-cost fallacy. The data analysis techniques used in this study were Item Factor Analysis (IFA) and graded response models. Respondents in this study were 217 students in Indonesia. The analysis results show that the 1-factor model of resistance to sunk costs fits the data. At the item level, based on the IFA, it was found that two out of ten items were found to be misfits. The results of the GRM analysis on the remaining eight items indicate that the items fit the GRM. However, the distinguishing power and reliability were found to be low. It can be concluded that this preliminary study provides a variety of important information as suggestions for improvement for our modified scale. Several theoretical and methodological implications are discussed.

**Keywords:** item response theory; situational judgment test; sunk cost; validation

Imagine the following situation, taken from de Bruin et al. (2007) experiment. You are buying a gold ring on layaway for someone special. It costs \$200 and you have already paid \$100 on it to Store A and still owe another \$100. However, one day you see in the paper that a new jewelry store is selling the same ring for only \$90 on a special sale, and you can still pay for it using layaway. The new store, Store B, is across the street from Store A. If you decide to get the ring from the new store, you will not be able to get your money back from the old store, but you would save \$10 overall. In that situation, would you be more likely to continue paying at Store A or buy from Store B?

The situation above describes the decision-making that centers on sunk-cost, which is a cost or investment that has been incurred and cannot be returned. Rationally, the decisions that should be taken are very clear, namely only incremental costs and benefits should affect decisions (Thaler, 1980; Tversky & Kahneman, 1981). Ordering a ring at Shop B is more profitable, but in a situation like this, many people still choose to buy at Shop A. An irrational decision like this is known as the sunk-cost effect or the sunk-cost fallacy (Arkes & Blumer, 1985; Dijkstra & Hong, 2019; Jarmolowicz et al., 2016)

<sup>\*</sup>Address for correspondence: r.hidayat@ugm.ac.id

or Concorde effects in the research literature on animal subjects (Arkes & Ayton, 1999).

The sunk-cost fallacy is a pervasive anomaly in decision-making (Dijkstra & Hong, 2019; Roth et al., 2014; J. Wang & Keil, 2007). The effect of sunk-cost has been observed consistently in experiments with prior investment in the form of money (Arkes & Blumer, 1985; Garland & Newport, 1991; Rosenbaum & Lamort, 1992; Tversky & Kahneman, 1981), and behavioral investment, such as time and effort, (Coleman, 2010; Macaskill & Hackenberg, 2012; Marcus Cunha & Caldieraro, 2009; Soman & Cheema, 2001). Most studies have been conducted with adult participants, but the sunk-cost fallacy has also been observed in studies with young participants (Whitely & Dawis, 1974), and even in studies with animal subjects (de Magalhães et al., 2011; Macaskill & Hackenberg, 2012; Magalhaes & White, 2016), although other researchers argue that the sunk-cost effect is not found in animal behavior (Arkes & Ayton, 1999; Yáñez et al., 2017). The sunk-cost fallacy is also found in the context of interpersonal relationships (Ostrom et al., 2018; Rego et al., 2018), moral judgments (Meyers et al., 2020), entrepreneurial settings (McCarthy et al., 1993), corporate management (Chung & Cheng, 2018), the agricultural and food industries (Boland et al., 2014), is related to psychotherapy and mental health issues (Leahy, 2000; Schmitzer-Torbert, 2020), and is even associated with the collapse of ancient societies (Janssen et al., 2003).

#### *Decision-Making Experiments Involving Sunk-Cost*

Most of the research on the sunk-cost fallacy has been conducted using experimental methods. As an experimental treatment, participants were exposed to various forms of decision-making problems involving sunk costs. Experiments were conducted on a paper-and-pencil basis or via computer (Arkes & Blumer, 1985; Dijkstra & Hong, 2019; Garland & Newport, 1991; Tversky & Kahneman, 1981; Webley & Plaisier, 1998), in studies involving monetary incentives that are in line with the outcome of the decisions made by participants (Heath, 1995). Experiments were also carried out in the form of real commitments in the field, for example, the manipulation of discounts on cinema subscription packages on campus in the second experiment from Arkes and Blumer (1985), or using historical data (Q. A. W. Keefer, 2016, 2019). Various experiments have been conducted to study the factors that contribute to the sunk-cost fallacy, for example, the level of investment (Arkes & Ayton, 1999) and windfall gains (Soman & Cheema, 2001).

The basic assumption underlying these experimental studies is that the sunk-cost effect is a behavior caused by factors external to the individual, for example, the probability of success is related to the achievement of goals in making a prior investment (Arkes & Hutzler, 2000), as long as the funds involved in sunk-cost, whether regular income or windfall gains (Soman & Cheema, 2001). This is an assumption inherent in experimental research methodology (Shadish et al., 2004). Experimental manipulation is carried out to examine the internal factors and mechanisms of the psychological processes underlying the sunk-cost effect, such as mental budgeting which affects the de-escalation of commitment in a sunk-cost situation (Heath, 1995), and the rule of do not waste in decision-making (Arkes & Ayton, 1999; Zultan et al., 2010), personal responsibility (McCarthy et al., 1993), the negative emotions inherent in decision-making situations (Dijkstra & Hong, 2019), cognitive dissonance (Chung

& Cheng, 2018), neural mechanisms (Fujino et al., 2016; Zeng et al., 2013), and theoretical knowledge about the principles of economic decision-making (Rover et al., 2009). However, the assumption held is that individuals will act with the same mental and neural processes, and produce similar behavioral tendencies.

#### *Individual Differences in Vulnerability to The Sunk-Cost Fallacy*

The idea of individual differences in the sunk-cost fallacy, and in rationality in general, is interesting to explore. First, the sunk-cost fallacy was not observed in 100% of participants in the experiments (Ashraf et al., 2006; Jarmolowicz et al., 2016). Furthermore, there are differences in the propensities of sunk costs between settings and experimental manipulations, as well as between specific contexts of decision-making (Haita-Falah, 2017). This opens up possible interpretations regarding individual differences in vulnerability to the sunk-cost fallacy across decision-making situations.

The notion of a psychological construct in individual differences is supported by several studies based on different theoretical arguments. For example, Hafenbrack et al. (2013) found that the sunk-cost fallacy is related to a person's level of depression. Other studies found links between the sunk-cost fallacy and various psychological symptoms (Jarmolowicz et al., 2016), increased activity in lateral frontal and parietal cortices which are normally involved in risk-taking (Zeng et al., 2013), and emotional conditions and personality traits of anxiety (Dijkstra & Hong, 2019). Some other studies found that young people are more likely to engage in the sunk-cost fallacy compared to adults, especially when it involves money (Roth et al., 2014; Strough et al., 2008).

Several studies measure and demonstrate individual differences in the sunk-cost fallacy. de Bruin et al. (2007) measured individual differences in the sunk-cost fallacy as part of individual differences in adult decision-making competence. Haita-Falah (2017) examined the role of cognitive ability as a differentiating factor in susceptibility to the sunk-cost effect. Fujino et al. (2016) found that trait agreeableness and conscientiousness, with the mediation of insula activity in neural mechanisms, have an effect on individuals' susceptibility to sunk-cost effects. Kwak and Park (2011) suggested that alignment between individual focus regulators and problem characteristics increases vulnerability to sunk-cost effects.

For individual differences to be valid and reliably measured, an instrument for observing the sunk-cost fallacy that satisfies good psychometric criteria must first be developed. Considering that individual differences need to be observed in consistent behavior across different situations, the various hypothetical scenarios of decision-making situations involving the sunk-cost fallacy that are studied in the experiments (Arkes & Blumer, 1985; Dijkstra & Hong, 2019; Haita-Falah, 2017; Jarmolowicz et al., 2016; Moon, 2001) can be utilized. However, the measurement should be carried out in a non-experimental research design. One possible approach to measure the individual tendency toward the sunk-cost fallacy in these hypothetical situations is using a Situational Judgment Test (SJT).

*Situational Judgment Test for Psychological Measurement*

The SJT is a psychological construct measurement that involves conditions in the form of realistic or hypothetical scenarios and respondents are asked to assess in a certain way (Stemler & Sternberg, 2006). The SJT has been adapted into personality trait measurements (Olaru et al., 2019), prosocial implicit traits (Motowidlo et al., 2018), HEXACO personality traits (Oostrom et al., 2018), and perceptions of interpersonal behavior (Golubovich et al., 2016). In addition, various studies discuss the potential for adapting SJT to future research developments that are commonly termed construct-driven SJT (Campion et al., 2014; Guenole et al., 2017; Lievens, 2017).

Certainly, to the best of our knowledge, the construct-driven SJT can be adapted into the process of measuring resistance to sunk cost, which until now has not been found in the literature. The arrangement of the SJT in the measurement of the sunk-cost construct is possible, because since the beginning of the development of experimental-based theory about sunk-cost (Arkes & Blumer, 1985), the item's wording can be modified to have an ordinal scale response option (for example, a rating scale) and consequently, it can produce information about the trait level (level of resistance to sunk-cost) that is bipolar (a line from left to right) where the extreme left (low score) shows a low trait, while the extreme right (high score) shows a high trait.

In addition, from a methodological perspective, there are criticisms of SJT, such as a lack of supporting evidence about the factor structure of the measured constructs and advanced psychometric characteristics (Guenole et al., 2017). To date, the construct-driven SJT is considered different from the traditional SJT because it involves unidimensional response options so that it can describe the trait "level" of the respondent (Lievens, 2017). In parallel, the application of advanced psychometric methodologies such as item response theory (IRT) known as the 'latent trait theory' can be used to produce this information (detailed trait level and item level statistics), which is a recent topic in the development of studies on SJT data analysis (Ron, 2019). However, no studies have been conducted so far investigating the psychometric properties of the SJT of resistance to sunk-cost using the modern test theory. Thus, the application of IRT will provide a novel aspect in terms of the SJT resistance to sunk-cost data analysis methodology, where this method has long been termed item analysis which takes individual differences into account (Rasch, 1966).

Therefore, this study was conducted with the aim of (1) examining the dimensionality of the sunk-cost fallacy test with the SJT format; (2) evaluating the psychometric properties of the SJT using modern test theory; and (3) evaluating the weaknesses in terms of substance and psychometrics as suggestions for potential scalar refinement. This study will be the first contribution of modern test theory to assessing psychometric properties of the measurement of resistance to sunk-cost, specifically in the form of situational judgment tests.

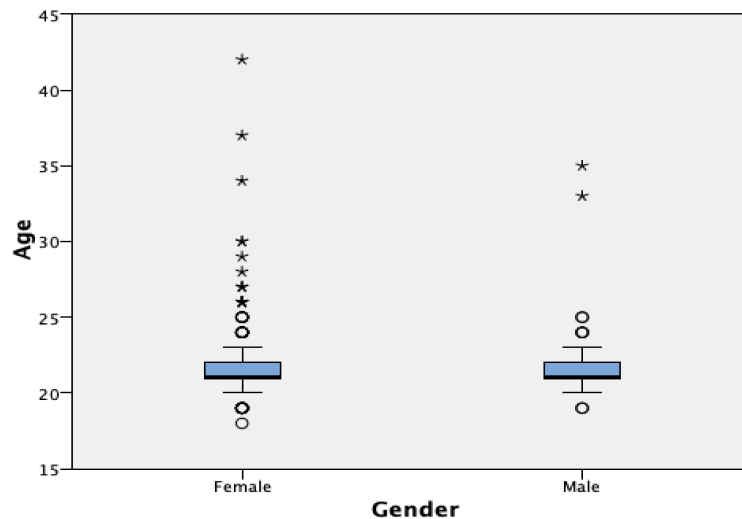
## Methods

### *Participants*

Participants were 217 students (60 male [27.6%], 157 female [72.4%]) at Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, with an age range of 18–42 years (mean age = 21.91,  $SD = 3.01$ ). The variation in respondents' ages for each gender can be seen in Figure 1. The purposive sampling method was used for data collection. The consideration in using this sampling technique was due to the limitations in terms of time to create a sample frame that contains data on active students, so purposive sampling was one of the non-probabilistic sampling techniques that were possible to use.

**Figure 1**

*Plot of Respondents' Ages Against Gender*



### *Instruments*

In this study, the Resistance to Sunk-Cost measurement instrument (RtSC) was adapted from the resistance to sunk costs measurement component, which is one of the aspects of the Adult Decision-Making Competence instrument (de Bruin et al., 2007). This instrument has been used in several other studies, including that by Dijkstra and Hong (2019). All 10 points of resistance to sunk-cost scenarios with a rating scale of 1–6 ranging from 1 (most likely to choose [the sunk-cost option]) to 6 (most likely to choose [the normatively correct option]) were also adapted in this study.

To improve relevance with social, economic, and cultural factors taken into in the Indonesian samples, several items were modified, namely, item 3 (from setting designing costumes for a Halloween party to writing an essay), item 6 (from preparing a welcoming speech at a friend's wedding party to make a surprise gift for a friend's birthday), item 10 (from the situation of painting the room, to using shoes purchased from their own savings). In addition, we also consider the characteristics of the student sample in the modification process of the content, so we also test whether changes to

content can function well in a sample of Indonesian students.

#### *Data Analytical Procedures*

The RtSC data were fitted with the Graded Response Model (GRM), (Samejima, 1969), one of the most commonly used IRT models for polytomous items. The GRM is appropriate to use when dealing with ordered categories on a rating scale (Fraley et al., 2000). The GRM has a slope parameter and  $n-1$  threshold parameters for each item, where  $n$  is the number of response categories. The slope parameter measures item discrimination: that is, how well the item differentiates between higher versus lower levels of resistance to sunk-cost. Threshold parameters measure item difficulty; that is, the ease versus the difficulty of endorsing different response options for an item (Pilkonis et al., 2013). There are two basic assumptions of GRM application, namely: unidimensionality, which assumes that only one latent construct is measured by a set of items in an instrument and local independence, that a person's response to an item does not depend on the person's response to other items in an instrument (Mair, 2018; Whitely & Dawis, 1974).

For the assessment of dimensionality, an Item Factor Analysis (IFA) (Wirth & Edwards, 2007) was carried out. In the IFA, the model fit statistics of a 1-factor model including chi-square statistics, root mean square error of approximation (RMSEA), comparative fit index (CFI), and the standardized root mean square residual (SRMR) were investigated. The RMSEA should be less than 0.05 for an adequate fit; while the CFI should be greater than 0.90 and the SRMR should be less than 0.08 for an acceptable fit (J. Wang & Wang, 2019). For the assessment of local independence, the Q3 statistics (Yen, 1984) were inspected. A value of Q3 statistics greater than 0.20 indicates a local dependence between a pair of items (Christensen et al., 2016).

After the two assumptions of GRM had been met, the accuracy of the data against the model was evaluated using a fit index at both the overall model level (global fit) and the item level. The global fit indices used were  $C_2^2$  and other fit indices that are rooted in factor analysis such as RMSEA, SRMR and CFI. If the value of  $C_2^2$  was not significant, then the hypothesized model was fitted to the data (Cai & Monroe, 2014). In parallel, the RMSEA, SRMR, and CFI values had the same acceptance criteria as factor analysis which can be used in the IRT context (Cook et al., 2009).

After obtaining evidence of global fit on the instrument level, the item level fit statistics can be examined using the  $S^{-2}$  statistics. The item fitted to the GRM if the values of  $S^{-2}$  were not significant (Kang & Chen, 2008). In this study, the GRM model is estimated using the marginal maximum likelihood method which is implemented in the 'MIRT (Multidimensional Item Response Theory)' package (Chalmers, 2012) implemented in the RStudio program. In parallel, the IFA was carried out using the Mplus 8.5 program using the weighted least square with mean and variance adjusted estimator.

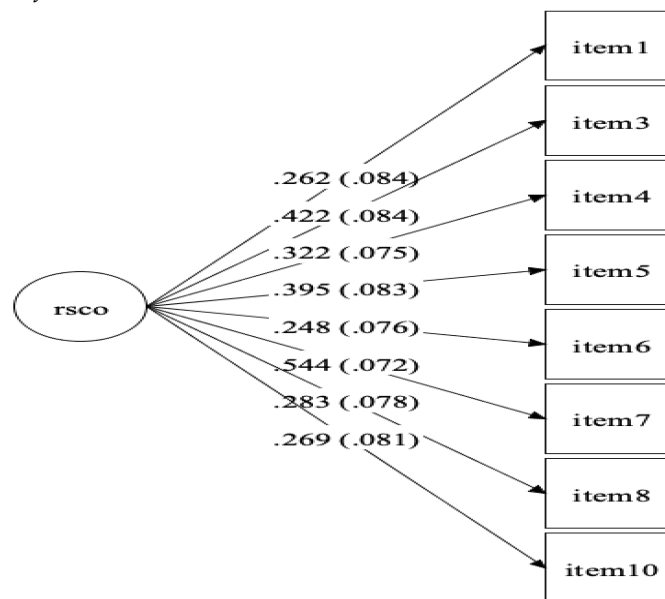
## Results

### *Unidimensionality: Item Factor Analysis*

The 1-factor IFA result shows evidence that the model does not fit to the data [ $\chi^2 = 83.891$ ,  $df = 35$ ,  $p = 0.000$ ,  $RMSEA = 0.080$  (90% C.I. = 0.058, 0.102),  $SRMR = 0.057$ ,  $CFI = 0.606$ ]. It was also found that the factor loading of two items (item 2 and item 9) was not statistically significant and also had a negative and very low factor loading. After both items were excluded from the analysis, the 1-factor IFA model was carried out on the remaining 8 items.

The results of the analysis show that the model fit to the data [ $\chi^2 = 30.128$ ,  $df = 20$ ,  $p = 0.068$ ,  $RMSEA = 0.048$  (90% CI = 0.000, 0.082),  $SRMR = 0.037$ ,  $CFI = 0.896$ ], means that the exclusion of two items (item 2 and item 9) from the analysis contributed to improving the fit of the model to the data. All items have a statistically significant factor loading in a positive direction in the range of 0.248 to 0.544. These findings indicate that the 1-factor model is a fit model in describing the resistance to sunk-cost construct and supports the use of a unidimensional GRM application. The graphical representation of the model can be seen in Figure 2.

**Figure 2**  
A One-Factor IFA Model of Resistance to Sunk-Cost Measure



### *Local Independence*

The local independence (LI) assumption test was conducted by inspecting the residual correlation between item pairs using the  $Q_3$  statistics (see Table 1). Based on the results of the analysis, no violation of the local independence assumption was found. None of the item pairs had any positive residual correlation with a value greater than 0.20, where the item pair that has the highest positive residual

correlation is item 8 and item 10 (residual correlation = 0.206). These findings support the application of GRM in this study.

**Table 1**  
*Residual Correlation Matrix for LI Testing*

Item	1	3	4	5	6	7	8	10
1	—							
3	-0.011	—						
4	-0.092	-0.044	—					
5	-0.025	-0.110	-0.126	—				
6	0.078	-0.127	-0.025	-0.049	—			
7	-0.178	-0.177	-0.172	-0.177	-0.147	—		
8	-0.043	-0.206	0.007	-0.072	-0.108	-0.046	—	
10	-0.037	-0.149	0.039	-0.189	-0.027	-0.098	0.103	—

*Item Measure, Fit Statistics and Item Characteristic Curves (ICC)*

The results of the GRM analysis of 8 items showed the acceptable global fit indices based on the predefined criteria ( $C2 = 22.956$ ,  $df = 20$ ,  $p = 0.290$ ,  $RMSEA = 0.026$  (90%  $CI = 0.000, 0.066$ ),  $CFI = 0.944$ ,  $SRMR = 0.062$ ). These findings indicate that the use of a unidimensional GRM in this study has been supported, in line with the findings of the IFA results. After GRM was found to fit to the data, the interpretation of the item level can be made. Item parameters estimated through the GRM can be seen in Table 2.

**Table 2**  
*Item Parameters and Fit Statistics for Each Item of The RtSC Measure*

Item	a	Threshold					Fit statistics		
		b1	b2	b3	b4	b5	$S - \chi^2$	df	Prob.
Item 1	0.452	-1.693	0.148	0.706	1.670	3.575	58.724	56	0.376
Item 3	0.926	-2.671	-1.631	-0.504	0.512	1.973	46.334	54	0.761
Item 4	0.620	-1.368	0.276	1.330	2.451	3.881	38.486	47	0.807
Item 5	0.809	-2.215	-1.361	-0.573	-0.265	1.134	52.798	50	0.366
Item 6	0.572	-2.506	-0.679	0.400	1.299	3.496	57.252	58	0.503
Item 7	1.191	-2.504	-1.708	-0.880	0.071	1.311	75.204	57	0.053
Item 8	0.512	-1.906	0.180	2.294	3.212	5.261	47.800	55	0.744
Item 10	0.515	-3.628	-2.240	-0.727	0.462	2.446	80.221	70	0.189

Table 2 contains information about the slope (item discrimination), threshold (item difficulty) and fit statistics of each item. It was found that the discriminating power of all items had a positive direction indicating that the items could function properly in distinguishing people with high trait tendencies from people with low traits. However, it was found that the distinguishing power of items tends to be low because it is in the range of 0.452 to 1.191. These findings will be discussed further.



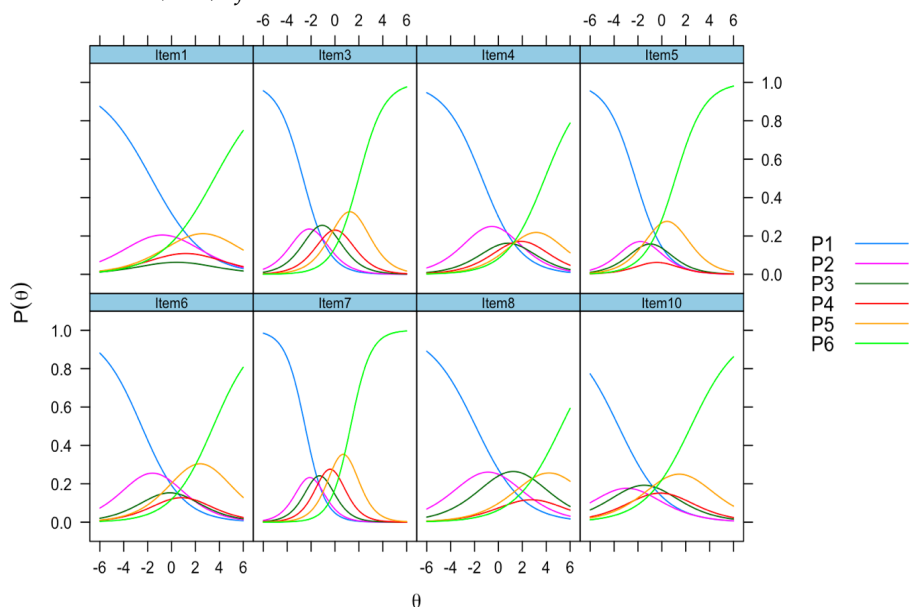
Various theories reveal that the “good” item discrimination is in the range of 0.8 to 2.5 (de Ayala, 2009), whereas in classical test theory and IRT perspectives, the higher item discrimination value shows better psychometric characteristics (Fraley et al., 2000). In this study, although the discriminating power is low, taking into account that there is no negative item discrimination, low item discrimination is important information for researchers to compile a stimulus that can distinguish between people with high resistance to sunk costs and those with low resistance.

In addition, it can be seen that all threshold values were ordered from lowest to highest and this applies to all items, indicating that no specific category had zero endorsements. This shows the level of the respondent’s sunk cost in choosing each response category. Based on the threshold magnitude, it can be seen that items 5 and 7 tend to get the highest response more easily, while item 8 has the most difficulty in getting the highest response.

After it was found that the entire model was fit for GRM, the information in Table 2 shows the accuracy of the model at the item level as seen based on the  $S^{-2}$  statistical test. Based on the analysis, it was found that all items fit the GRM. This shows that the use of GRM is appropriate in describing the construct measuring resistance to sunk cost.

The following information in Figure 3 shows the ICC for each item which is a visual representation of the characteristics of the item. Two items show a high peak, namely item 3 and item 7, which indicates that these two items have a high item information function compared to other items. In addition, it can be seen that in items 1 and 6, category 4 has very little chance of being selected.

**Figure 3**  
*Item Characteristic Curves (ICC) of Resistance to The Sunk-Cost Measure*

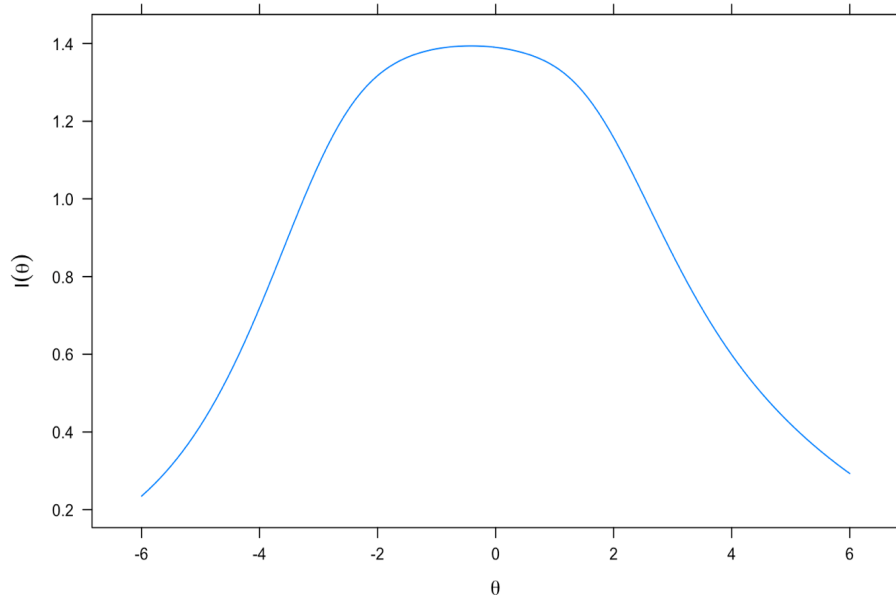


*Test Information Function (TIF) and Marginal Reliabilities*

The use of GRM also produces an estimate of the Total Information Function (TIF). The TIF provides information on the estimated information function test for each level of resistance to sunk cost (see Figure 3). The x-axis of the graph shows the level of resistance to the sunk cost of the respondent, while the y-axis of the graph shows the total information value.

As can be seen in Figure 3, the peak of the TIF is at the resistance level to sunk-cost 0.40 logit. This shows that this instrument will provide maximum information when used to assess persons with a low resistance to sunk costs. At the peak of the TIF, the standard error of the test is at the lowest value. However, the marginal reliability of the test score was found to be 0.575, which means that the internal consistency of the test score is quite low. This is closely related to the low item factor loading found in the IFA results. These findings will be discussed further.

**Figure 4**  
*Test Information Function (TIF) of Resistance to Sunk-Cost Measure*



**Discussion**

This preliminary study aimed to evaluate the psychometric characteristics of the measurement instrument for the resistance to the sunk-cost construct which was adapted and modified from one of the A-DMC components (de Bruin et al., 2007). The instrument was arranged in the form of an SJT in which IRT analysis was carried out to evaluate psychometric properties, which was also complemented by the use of IFA to test the construct validity of the instrument. Based on the IFA results it was found that the 1-factor model fitted to the data. However, even though the overall model fitted, at the item level, two items were found to have a negative factor charge, namely, item 2 and item 9. This means

that the two items functioned in a way that was opposite to what had been hypothesized.

In item 2, we found limitations in terms of item content, where the hypothetical scenario presented had language that the respondents could not understand. However, these findings only emerged after the data collection process was completed. Additionally, item 9 was found to have a very low and negative factor loading. It is clear that item 9 measures other constructs as it differs in terms of content from other items. The difference in item 9 was about the form of a hypothetical scenario that contains an assessment of the decisions taken by others, while the other 9 items contain an assessment of the decisions taken by oneself. The IFA findings support and provide information that there are problems in item 2 and item 9.

This information is very important as an initial reference for improving the instrument. We got information that in the process of scaling up, a hypothetical scenario can be constructed that contains “making decisions for oneself” and “making decisions for others.” Although many studies compare individual versus group decision-making (Davis & Toseland, 1987; Payne & Wood, 2002), we had difficulty finding studies that focus more on comparing stimulus for ‘individuals’ with ‘making decisions for others’ in the context of sunk-cost in particular. The findings of this preliminary study generate an idea for further exploration of this, considering that in the SJT, constraints on the constructs measured are more complex (Schmitt & Chan, 2006), so an in-depth analysis of content needs to be carried out which has already been recommended by previous studies (Campion et al., 2014).

In the next stage, the eight remaining items were analyzed by Polytomous IRT using GRM. In line with the IFA findings, the unidimensional model of GRM fits the data and all items are proven to fit the model. The consistency of these findings cannot be separated from the view that IFA and GRM are equivalent models (Kamata & Bauer, 2008). In addition, we found that there was no violation of the LI assumption between pairs of items. Theoretically, SJTs which contain a hypothetical scenario are independent from one another, which is the advantage of an SJT (Cabrera & Nguyen, 2001), this advantage certainly supports the assumption of LI from IRT as seen from our analysis results.

Based on the results of the GRM analysis, it was found that the discrimination power of items was classified as low (ranging from 0.452 to 1.191), with item 7 ( $a = 1.191$ ) having the highest discrimination power. This is closely related to the low factor loading in IFA (ranging from 0.248 to 0.544) (Fraleley et al., 2000). Referring to the acceptable discriminating power criteria of 0.8 to 2.5 (de Ayala, 2009), of the eight items, three have discriminating power in that range (item 3  $a = 0.926$ ; item 5  $a = 0.809$ ; item 7  $a = 1.191$ ). From this information, we have important insights to examine the content of the item and suggest further development, regarding what forms of hypothetical scenarios have good discrimination power to be able to distinguish people with a high resistance to sunk costs compared to those with a low resistance to sunk-cost.

In the end, the IFA factor loading and the low discrimination power of some items based on GRM analysis have an impact on the internal consistency (reliability) of the RtSC scores, which was found to be the marginal reliability of this instrument at 0.525. The low reliability of SJT and its causative factors has been discussed for a long time (Campion et al., 2014; Guenole et al., 2017). Given that an SJT will function optimally to measure very specific constructs (Schmitt & Chan, 2006), we

realize that although resistance to sunk-cost is a very specific construct, some aspects have not been considered in the cultural adaptation process carried out. The aspect in question is related to the theoretical framework of sunk-cost which involves emotional aspects (Dijkstra & Hong, 2019), the form of investment (Jarmolowicz et al., 2016; Marcus Cunha & Caldieraro, 2009), which will be taken into account in the future development of SJT-RtSC.

We suspect that although the factor loading of all items was significant and the 1-factor model fitted to the data, the low factor loading will be resolved when the context of measuring the resistance to sunk-cost is addressed by a tightening in the grouping of the hypothetical scenarios. In addition, when hypothetical scenarios are grouped before data collection, various models can be compared which represent the construct of resistance to sunk costs. This is one of the limitations of our preliminary study, however, we arrived at this information based on the results of this preliminary study.

Furthermore, it is important to note that this study is not without limitations, one of which is the selected samples. The respondents were chosen from only one university using a non-probabilistic sampling technique. Thus, the findings of this study cannot be over-generalized for different samples' characteristics and the generalization is limited to our samples only or another sample with similar characteristics (i.e., university students). Second, this cross-sectional validation study only provides evidence of validity from one point of view, which is construct validity. Although the approach was acceptable, another type of evidence of validity such as convergent and consequential validity testing would generate an interesting assessment of psychometric properties of resistance to sunk costs measures and also follows the recommendations of the Standards for Educational and Psychological Testing from APA, AERA & NCME (2014) which recommend the reporting of five instances of evidence of validity. Despite these limitations, our findings can serve as a reference for future research in assessing the relationship between resistance to sunk-cost with other factors or variables.

## Conclusion

This paper aims to report the results of the evaluation of the psychometric properties of the instrument for measuring individual differences in the sunk-cost fallacy with the SJT format and identify weaknesses in terms of substance and psychometrics as suggestions for potential scale refinement. This study shows that the psychometric properties of adaptation instruments are generally acceptable. Our findings are the first empirical evidence of an adequate internal structure of an Indonesian-adapted version of the SJT-RtSC. Based on preliminary validation, we found several aspects that can be improved on both in terms of theoretical and methodological frameworks. In the end, this adaptation instrument contributes to explaining the combination of SJT and modern test theory in a very limited sunk-cost measurement instrument. Further studies are expected to overcome our limitations so that an instrument can be used to measure resistance to sunk-cost comprehensively and can be widely used.

### *Recommendation*

Eight of the 10 items adapted from the A-DMC (de Bruin et al., 2007) can be considered for application in the measurement of individual differences in sunk-cost fallacies. However, further research is needed to test the measurement of sunk-cost fallacies in functional areas such as in organizational and business decision making. In light of the results from this research, the use of the SJT format in these research are encouraged.

## **Declaration**

### *Acknowledgement*

The researchers would like to thank all participants for the voluntary participations in this research. We also would like to than all parties in the Faculty of Psychology Universitas Gadjah Mada for the support throughout this research.

### *Conflic of Interests*

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### *Funding*

The author's received no financial support for the research, authorship, and/or publication of this article.

### *Author Contribution*

RH conducted instrument adaptation, designed the study, and lead in the data collection. MDKP designed and conducted the data analysis. Both authors wrote and approved the final version of the manuscript.

### *Orcid ID*

Rahmat Hidayat  <https://orcid.org/0000-0002-1323-2914>

Muhammad Dwirifqi Kharisma Putra  <http://orcid.org/0000-0002-9383-7904>

## **References**

- Arkes, H. R., & Ayton, P. (1999). The sunk cost and concorde effects: Are humans less rational than lower animals? *Psychological Bulletin*, 125(5), 591–600. <https://doi.org/10.1037/0033-2909.125.5.591>
- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1), 124–140. [https://doi.org/10.1016/0749-5978\(85\)90049-4](https://doi.org/10.1016/0749-5978(85)90049-4)

- Arkes, H. R., & Hutzler, L. (2000). The role of probability of success estimates in the sunk cost effect. *Journal of Behavioral Decision Making*, 13(3), 295–306. [https://doi.org/10.1002/1099-0771\(200007/09\)13:3<295::aid-bdm353>3.0.co;2-6](https://doi.org/10.1002/1099-0771(200007/09)13:3<295::aid-bdm353>3.0.co;2-6)
- Ashraf, N., Karlan, D., & Yin, W. (2006). Household decision making and savings impacts: Further evidence from a commitment savings product in the philippines. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.912771>
- Boland, M. A., Crespi, J. M., & Turner, T. M. (2014). Measuring sunk costs in agricultural and food industry assets: Why some assets sell below appraisal. *Journal of Agricultural & Food Industrial Organization*, 12(1). <https://doi.org/10.1515/jafio-2014-0009>
- Cabrera, M. A. M., & Nguyen, N. T. (2001). Situational judgment tests: A Review of practice and constructs assessed. *International Journal of Selection and Assessment*, 9(1&2), 103–113. <https://doi.org/10.1111/1468-2389.00167>
- Cai, L., & Monroe, S. (2014). *A New statistic for evaluating item response theory models for ordinal data*. University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). <https://eric.ed.gov/?id=ED555726>
- Campion, M. C., Ployhart, R. E., & MacKenzie, W. I. (2014). The state of research on situational judgment tests: A content analysis and directions for future research. *Human Performance*, 27(4), 283–310. <https://doi.org/10.1080/08959285.2014.929693>
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6). <https://doi.org/10.18637/jss.v048.i06>
- Christensen, K. B., Makransky, G., & Horton, M. (2016). Critical values for yen's Q3: Identification of local dependence in the rasch model using residual correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>
- Chung, S.-H., & Cheng, K.-C. (2018). How does cognitive dissonance influence the sunk cost effect? *Psychology Research and Behavior Management*, Volume 11, 37–45. <https://doi.org/10.2147/prbm.s150494>
- Coleman, M. D. (2010). Sunk cost and commitment to medical treatment. *Current Psychology*, 29(2), 121–134. <https://doi.org/10.1007/s12144-010-9077-7>
- Cook, K. F., Kallen, M. A., & Amtmann, D. (2009). Having a fit: Impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Quality of Life Research*, 18(4), 447–460. <https://doi.org/10.1007/s11136-009-9464-4>
- Davis, L., & Toseland, R. (1987). Group versus individual decision making. *Social Work With Groups*, 10(2), 95–110. [https://doi.org/10.1300/j009v10n02\\_09](https://doi.org/10.1300/j009v10n02_09)
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.
- de Bruin, W. B., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, 92(5), 938–956. <https://doi.org/10.1037/0022-3514.92.5.938>

- de Magalhães, J. P., Wuttke, D., Wood, S. H., Plank, M., & Vora, C. (2011). Genome-environment interactions that modulate aging: Powerful targets for drug discovery. *Pharmacological Reviews*, 64(1), 88–101. <https://doi.org/10.1124/pr.110.004499>
- Dijkstra, K. A., & Hong, Y.-y. (2019). The feeling of throwing good money after bad: The role of affective reaction in the sunk-cost fallacy. *Plos One*, 14(1), e0209900. <https://doi.org/10.1371/journal.pone.0209900>
- Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, 78(2), 350–365. <https://doi.org/10.1037/0022-3514.78.2.350>
- Fujino, Y., Amimoto, K., Sugimoto, S., Fukata, K., Inoue, M., Takahashi, H., & Makita, S. (2016). Prone positioning reduces severe pushing behavior: Three case studies. *Journal of Physical Therapy Science*, 28(9), 2690–2693. <https://doi.org/10.1589/jpts.28.2690>
- Garland, H., & Newport, S. (1991). Effects of absolute and relative sunk costs on the decision to persist with a course of action. *Organizational Behavior and Human Decision Processes*, 48(1), 55–69. [https://doi.org/10.1016/0749-5978\(91\)90005-e](https://doi.org/10.1016/0749-5978(91)90005-e)
- Golubovich, J., Seybert, J., Martin-Raugh, M., Naemi, B., Vega, R. P., & Roberts, R. D. (2016). Assessing perceptions of interpersonal behavior with a video-based situational judgment test. *International Journal of Testing*, 17(3), 191–209. <https://doi.org/10.1080/15305058.2016.1194275>
- Guenole, N., Chernyshenko, O. S., & Weekly, J. (2017). On designing construct driven situational judgment tests: Some preliminary recommendations. *International Journal of Testing*, 17(3), 234–252. <https://doi.org/10.1080/15305058.2017.1297817>
- Hafenbrack, A. C., Kinias, Z., & Barsade, S. G. (2013). Debiasing the mind through meditation. *Psychological Science*, 25(2), 369–376. <https://doi.org/10.1177/0956797613503853>
- Haita-Falah, C. (2017). Sunk-cost fallacy and cognitive ability in individual decision-making. *Journal of Economic Psychology*, 58, 44–59. <https://doi.org/10.1016/j.joep.2016.12.001>
- Heath, C. (1995). Escalation and de-escalation of commitment in response to sunk costs: The role of budgeting in mental accounting. *Organizational Behavior and Human Decision Processes*, 62(1), 38–54. <https://doi.org/10.1006/obhd.1995.1029>
- Janssen, I., Krabbendam, L., Jolles, J., & van Os, J. (2003). Alterations in theory of mind in patients with schizophrenia and non-psychotic relatives. *Acta psychiatrica Scandinavica*, 108, 110–7. <https://doi.org/10.1034/j.1600-0447.2003.00092.x>
- Jarmolowicz, D. P., Bickel, W. K., Sofis, M. J., Hatz, L. E., & Mueller, E. T. (2016). Sunk costs, psychological symptomology, and help seeking. *SpringerPlus*, 5(1). <https://doi.org/10.1186/s40064-016-3402-z>
- Kamata, A., & Bauer, D. J. (2008). A Note on the relation between factor analytic and item response theory models. *Structural Equation Modeling: A Multidisciplinary Journal*, 15(1), 136–153. <https://doi.org/10.1080/10705510701758406>

- Kang, T., & Chen, T. T. (2008). Performance of the generalized S-X2 item fit index for polytomous IRT Models. *Journal of Educational Measurement*, 45(4), 391–406. <https://doi.org/10.1111/j.1745-3984.2008.00071.x>
- Keefer, Q. A. W. (2016). The sunk-cost fallacy in the national football league. *Journal of Sports Economics*, 18(3), 282–297. <https://doi.org/10.1177/1527002515574515>
- Keefer, Q. A. (2019). Decision-maker beliefs and the sunk-cost fallacy: Major league baseball's final-offer salary arbitration and utilization. *Journal of Economic Psychology*, 75, 102080. <https://doi.org/10.1016/j.joep.2018.06.002>
- Kwak, J., & Park, J. (2011). Effects of a regulatory match in sunk-cost effects: A mediating role of anticipated regret. *Marketing Letters*, 23(1), 209–222. <https://doi.org/10.1007/s11002-011-9148-z>
- Leahy, R. L. (2000). Sunk costs and resistance to change. *Journal of Cognitive Psychotherapy*, 14(4), 355–371. <https://doi.org/10.1891/0889-8391.14.4.355>
- Lievens, F. (2017). Construct-driven (SJTs): Toward an agenda for future research. *International Journal of Testing*, 17(3), 269–276. <https://doi.org/10.1080/15305058.2017.1309857>
- Macaskill, A. C., & Hackenberg, T. D. (2012). The sunk cost effect with pigeons: Some determinants of decisions about persistence. *Journal of The Experimental Analysis of Behavior*, 97(1), 85–100. <https://doi.org/10.1901/jeab.2012.97-85>
- Magalhaes, P., & White, K. G. (2016). The sunk cost effect across species: A review of persistence in a course of action due to prior investment. *Journal of The Experimental Analysis of Behavior*, 105(3), 339–361. <https://doi.org/10.1002/jeab.202>
- Mair, P. (2018). *Modern psychometrics with R*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-93177-7>
- Marcus Cunha, J., & Caldieraro, F. (2009). Sunk-cost effects on purely behavioral investments. *Cognitive Science*, 33(1), 105–113. <https://doi.org/10.1111/j.1551-6709.2008.01005.x>
- McCarthy, A. M., Schoorman, F., & Cooper, A. C. (1993). Reinvestment decisions by entrepreneurs: Rational decision-making or escalation of commitment? *Journal of Business Venturing*, 8(1), 9–24. [https://doi.org/10.1016/0883-9026\(93\)90008-s](https://doi.org/10.1016/0883-9026(93)90008-s)
- Meyers, E. A., Walker, A. C., Fugelsang, J. A., & Koehler, D. J. (2020). Reducing the number of non-naïve participants in Mechanical Turk samples. *Methods in Psychology*, 3, 100032. <https://doi.org/10.1016/j.metip.2020.100032>
- Moon, H. (2001). Looking forward and looking back: Integrating completion and sunk-cost effects within an escalation-of-commitment progress decision. *Journal of Applied Psychology*, 86(1), 104–113. <https://doi.org/10.1037/0021-9010.86.1.104>
- Motowidlo, S. J., Lievens, F., & Ghosh, K. (2018). Prosocial implicit trait policies underlie performance on different situational judgment tests with interpersonal content. *Human Performance*, 31(4), 238–254. <https://doi.org/10.1080/08959285.2018.1523909>
- Olaru, G., Burrus, J., MacCann, C., Zaromb, F. M., Wilhelm, O., & Roberts, R. D. (2019). Situational Judgment Tests as a method for measuring personality: Development and validity evidence



- for a test of dependability. *Plos One*, 14(2), e0211884. <https://doi.org/10.1371/journal.pone.0211884>
- Oostrom, J. K., de Vries, R. E., & de Wit, M. (2018). Development and validation of a (HEXACO) situational judgment test. *Human Performance*, 32(1), 1–29. <https://doi.org/10.1080/08959285.2018.1539856>
- Payne, J., & Wood, A. (2002). Individual decision making and group decision processes. *Journal of Psychology and Financial Markets*, 3(2), 94–101. [https://doi.org/10.1207/s15327760jpfm0302\\_04](https://doi.org/10.1207/s15327760jpfm0302_04)
- Pilkonis, P. A., Kim, Y., Yu, L., & Morse, J. Q. (2013). Adult Attachment Ratings (AAR): An item response theory analysis. *Journal of Personality Assessment*, 96(4), 417–425. <https://doi.org/10.1080/00223891.2013.832261>
- Rasch, G. (1966). An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology*, 19(1), 49–57. <https://doi.org/10.1111/j.2044-8317.1966.tb00354.x>
- Rego, A., Cunha, M., & Simpson, A. (2018). The perceived impact of leaders' humility on team effectiveness: An empirical study. *Journal of Business Ethics*, 148, 205–218. <https://doi.org/10.1007/s10551-015-3008-3>
- Ron, T. H. (2019). Bringing situational judgement tests to the 21st century: Scoring of situational judgement tests using item response theory.
- Rosenbaum, D. I., & Lamort, F. (1992). Entry, barriers, exit, and sunk costs: An analysis. *Applied Economics*, 24(3), 297–304. <https://doi.org/10.1080/00036849200000142>
- Roth, S., Robbert, T., & Straus, L. (2014). On the sunk-cost effect in economic decision-making: A Meta-analytic review. *Business Research*, 8(1), 99–138. <https://doi.org/10.1007/s40685-014-0014-8>
- Rover, S., Wuerges, A. F. E., Tomazzia, E. C., & Borba, J. A. (2009). Sunk costs effect: Does theoretical knowledge affects students' decision process? *Brazilian Business Review*, 6(3), 232–248. <https://doi.org/10.15728/bbr.2009.6.3.2>
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, 34(S1), 1–97. <https://doi.org/10.1007/bf03372160>
- Schmitt, N., & Chan, D. (2006). *Situational judgment tests: Method or construct?* Lawrence Erlbaum Associates Publishers.
- Schmitzer-Torbert, N. (2020). Mindfulness and decision making: Sunk costs or escalation of commitment? *Cognitive Processing*, 21(3), 391–402. <https://doi.org/10.1007/s10339-020-00978-4>
- Shadish, W., Cook, T., & Campbell, D. (2004). Quasi-experimental designs for generalized causal inference.
- Soman, D., & Cheema, A. (2001). The effect of windfall gains on the sunk-cost effect. *Marketing Letters*, 12(1), 51–62. <https://doi.org/10.1023/a:1008120003813>
- Stemler, S. E., & Sternberg, R. J. (2006). *Using situational judgment tests to measure practical intelligence.* Lawrence Erlbaum Associates Publishers.

- Strough, J., Mehta, C. M., McFall, J. P., & Schuller, K. L. (2008). Are older adults less subject to the sunk-cost fallacy than younger adults? *Psychological Science*, *19*(7), 650–652. <https://doi.org/10.1111/j.1467-9280.2008.02138.x>
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39–60. [https://doi.org/10.1016/0167-2681\(80\)90051-7](https://doi.org/10.1016/0167-2681(80)90051-7)
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- Wang, J., & Wang, X. (2019). *Structural equation modeling*. <https://doi.org/10.1002/9781119422730>
- Wang, J., & Keil, M. (2007). A meta-analysis comparing the sunk cost effect for IT and Non-IT projects. *Information Resources Management Journal*, *20*(3), 1–18. <https://doi.org/10.4018/irmj.2007070101>
- Webley, P., & Plaisier, Z. (1998). Mental accounting in childhood. *Citizenship, Social and Economics Education*, *3*(2), 55–64. <https://doi.org/10.2304/csee.1998.3.2.55>
- Whitely, S. E., & Dawis, R. V. (1974). The nature of objectivity with the rasch model. *Journal of Educational Measurement*, *11*(3), 163–178. <https://doi.org/10.1111/j.1745-3984.1974.tb00988.x>
- Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods*, *12*(1), 58–79. <https://doi.org/10.1037/1082-989x.12.1.58>
- Yáñez, N., Bouzas, A., & Orduña, V. (2017). Rats behave optimally in a sunk cost task. *Behavioural Processes*, *140*, 47–52. <https://doi.org/10.1016/j.beproc.2017.04.003>
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, *8*(2), 125–145. <https://doi.org/10.1177/014662168400800201>
- Zeng, J., Zhang, Q., Chen, C., Yu, R., & Gong, Q. (2013). An fMRI study on sunk cost effect. *Brain Research*, *1519*, 63–70. <https://doi.org/10.1016/j.brainres.2013.05.001>
- Zultan, R., Bar-Hillel, M., & Guy, N. (2010). When being wasteful appears better than feeling wasteful. *Judgment and Decision Making*, *5*, 489–496. <https://journal.sjdm.org/10/10919/jdm10919.html>