Methods in Psychology 5 (2021) 100081

Contents lists available at ScienceDirect

Methods in Psychology

journal homepage: www.sciencedirect.com/journal/methods-in-psychology

Can Likert scales predict choices? Testing the congruence between using Likert scale and comparative judgment on measuring attribution☆

Che Cheng[a], Keng-Ling Lay[a,**], Yung-Fong Hsu[a,*], Yi-Miau Tsai[b]

[a] Department of Psychology, National Taiwan University, Taiwan
[b] Institute for Social Research, University of Michigan, USA

ARTICLE INFO

Keywords:
Choice
Comparative judgment
Likert scale
Preference

ABSTRACT

How people make choices among alternatives are of interest in different areas of psychological research. One paradigm to answer this question is by applying Likert Scale (LS) to compare the agreements to different alternatives, and the respective LS scores are then transferred into rank order of preference. However, using LS to infer choices is somewhat debatable because the measurement format of LS was not designed for revealing psychological preference. In this article, we examined to what extent it is appropriate to use quantitative indicators derived from LS to infer choices, with which we used the Comparative Judgment (CJ) procedure to represent a direct measurement of choice decision to compare. A total of 929 adolescents reported their effort and ability attributions for academic failure and success using both LS and CJ. We found that while using LS is generally accurate in predicting results obtained via CJ, the percentage of people revealing different choices inferred from LS versus CJ was 14.7% and 12.1% for the success and failure scenarios, respectively, suggesting that inferring psychological preference from LS is not without risk, at least for this sample of adolescents from a culturally-Chinese society. Furthermore, the majority of participants displaying incongruent decisions of achievement attribution via LS and CJ showed equivalent LS scores between effort and ability attributions. A goodness-of-fit test was conducted (on a model motivated by the beta-binomial distribution) and successfully eliminated the possibility that the tied LS scores in effort and ability attributions actually represent participants' true psychological state.

1. Introduction

People face plenty of choices in everyday life. Theories to explain and predict how people make choices among alternatives have been explored in different areas of psychological research. For example, in the area of achievement motivation, researchers often investigate individuals' preferences of goal orientation (Ames and Archer, 1987; Elliott and Dweck, 1988), attribution (Nicholls, 1976; Weiner and Kukla, 1970; Weiner, 1985), and implicit theory of intelligence (Dweck, 1986; Dweck and Leggett, 1988). In these areas, one of the most widely adopted measurement tools is the so-called *Likert scale (LS)*.

The LS is a commonly used rating format to represent participants' levels of agreement to statements of interest. The ratings are integer-valued on a given range (e.g., from 1 [strongly disagree] to 7 [strongly agree]), and the final outcomes are typically obtained by taking the summation or average of scores of individual items (Likert, 1932). Since its procedures of scale construction and strategies for data analysis have been well-established, the LS paradigm has been widely adopted in different topics of psychological research. Choice-related topics are without exception. For instance, researchers in the area of achievement motivation are concerned about whether people pick effort or ability as the major attribution of their academic success/failure (Nicholls, 1976). In such types of studies, LS is applied to measure the agreement level of different alternatives (e.g., González et al., 2018;

☆ The SAS/R codes used for all analyses are available at https://osf.io/m3bxe/. The real data used for the current analysis was retrieved from the unpublished study conducted by Lay and Tsai (2005). We have no known conflict of interest to disclose. The present study was supported in part by the Ministry of Science and Technology of Taiwan (R.O.C.) Grants 109-2410-H-002-092-MY2 (Keng-Ling Lay, PI) and 109-2410-H-002-088-MY2 (Yung-Fong Hsu, PI). Further support to Keng-Ling Lay was provided by the Program for Promoting Academic Excellent of Universities, Ministry of Education of Taiwan (R.O.C.) Grant 89-H-FA01-2-4.

* Corresponding author. Department of Psychology, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 106319, Taiwan.
** Corresponding author.
E-mail addresses: kllay@ntu.edu.tw (K.-L. Lay), yfhsu@ntu.edu.tw (Y.-F. Hsu).

https://doi.org/10.1016/j.metip.2021.100081
Received 29 June 2020; Received in revised form 4 August 2021; Accepted 19 October 2021
Available online 25 October 2021

Marsh, 1984; Tsujimoto et al., 2018; Vuletich et al., 2018), and the respective LS scores are then transferred into rank order or priority (e.g., Chen et al., 2009). For example, to assess the factors to which one's academic success was attributed, Chen et al. (2009) used the statement "I am talented" as the LS item for ability and the statement "I worked hard" for effort. Furthermore, they assumed that participants would consider the attribution with the highest LS score as the most probable reason for their academic success. Plenty of examples can also be found in other areas of psychological research (e.g., Harmon-Jones et al., 2018; Rentfrow and Gosling, 2003; Wang et al., 2019).

However, applying LS to infer human choices and preferences is somewhat debatable because the measurement format of LS was not designed for revealing psychological preference *per se*. Moreover, the issues of reliability and validity of LS were mostly studied based on *inter*-individual correlations (Campbell and Fiske, 1959; Lord et al., 1968), and it is unclear whether one can use the results of such studies to infer *intra*-individual choice judgments.

How to accurately measure and represent human choices and preferences is a methodological issue yet to be studied thoroughly. In this article we examined whether it is appropriate to use quantitative indicators derived from the results of LS to infer choices among alternatives, with which we used the (binary) comparative judgment (CJ) procedure (Thurstone, 1927) to represent a direct measurement of choice decision to compare. In a CJ procedure, participants respond to each question based on their own preference for the two alternatives. While prior studies have found moderate to strong correlations between LS and the Thurstone scale constructed by CJ (Bartlett et al., 1960; Joubert et al., 2015; Likert, 1932), the correlational analyses do not suffice in providing information about the congruence of the intra-individual choices predicted by CJ and by LS. To this end, we systematically collected data of both LS and CJ from the same group of participants, and examined the congruence of the choices inferred by the two.

This article is organized as follows. We first briefly review the scoring methods for LS and CJ, and describe the common practice usually applied to aggregate the intra-individual data for the respective method. We then lay out a procedure for assessing the degree of congruence between the choices inferred from the two. Next, we implement the procedure by using real data concerning academic attributions, and particularly focus on analyzing the cases displaying incongruent responses to LS and CJ. We wrap up the article with a brief discussion of the methodological issues regarding model fitting for both LS and CJ data.

1.1. Congruence of choices inferred from LS and from CJ

For each pair of alternatives X and Y, we aim to examine whether the choice inferred by LS items is congruent with the choice inferred by CJ items. Let L_{iXk} (or L_{iYk}) be the LS score regarding the agreement on the statement of item X_k (or Y_k) by an individual i. Thus both item X_k and item Y_k describe the same condition of k just that L_{iXk} states the response of X and L_{iYk} states the response of Y. A corresponding CJ item (C_{iXYk}) requires participant i to choose between X and Y as which is more preferable under condition k.

A concrete example is the academic attribution inventory described in the Appendix. To illustrate, let effort (X) and ability (Y) be the two attributions, and subjects are presented with a scenario, such as "Imagine that you've got a good grade and performed excellently in an exam." After reading a particular scenario, they were then presented with three conditions (i.e., $k = 1, 2, 3$), such as "After the exam, how would you usually consider a situation like this from the bottom of your heart?" ($k = 1$). Each condition is followed by three questions regarding attribution. Take $k = 1$ as an example. For an individual i,

- L_{iX1} is the response to the LS item "I performed well because I had worked hard.";
- L_{iY1} is the response to the LS item "I performed well because I am good at it.";
- C_{iXY1} is the response to the CJ item "If you could choose only one reason, which of the above is more like the reason for your good performance? Working hard or being good at it?".

A common way of inferring participants' choices from their LS data is to assume that they would choose the alternative with a higher aggregated LS score. A tie is assigned if the scores of the two alternatives are the same (Dittrich et al., 2007). Specifically, let $L_{iX} = \sum_{k=1}^{K} L_{iXk}/K$ and $L_{iY} = \sum_{k=1}^{K} L_{iYk}/K$ be the respective average score of the individual i. Then the inferred choice between X and Y can be conveniently coded as follows:

$$L_{iXY} = \left\{ \begin{array}{|l|c|c|} \hline +1, & L_{iX} < L_{iY} & \text{Choosing } Y \\ 0, & L_{iX} = L_{iY} & \text{LS tie} \\ -1, & L_{iX} > L_{iY} & \text{Choosing } X. \\ \hline \end{array} \right. \tag{1}$$

We call this rule of inferring choices from LS the *discrepancy rule*.

In the case of CJ, researchers often adopt a so-called *majority rule* to infer the final choice according to responses to several CJ items. The majority rule states that a person's preference is decided based on the alternative chosen more often (Arrow, 2012). A tie happens when the frequency of choosing X is equal to the frequency of choosing Y. Let $C_{iXYk} = 1$ for choosing Y and $C_{iXYk} = -1$ for choosing X by participant i, then based on the majority rule of CJ, the choice between pair (X, Y) can be coded as follows:

$$C_{iXY} = \left\{ \begin{array}{|l|c|c|} \hline +1, & \sum_k C_{iXYk} > 0 & \text{Choosing } Y \\ 0, & \sum_k C_{iXYk} = 0 & \text{CJ tie} \\ -1, & \sum_k C_{iXYk} < 0 & \text{Choosing } X. \\ \hline \end{array} \right. \tag{2}$$

These notations are summarized in Table 1. Note that there is no CJ tie if K is an odd number.

For each individual, we can assess the congruence of choices inferred from the two paradigms of LS and CJ based on the concept of "concordance" and "discordance" in Kendall (1990). First, when there is neither LS tie nor CJ tie, it is obvious that the two choosing results are either concordant or discordant. If both inferred choices show no preference between the two alternatives, the two results are classified as being concordant. In the case that only one of the two inferred choices shows a tie, we label it as "undetermined." Therefore, using the notation in Equations (1) and (2), an individual i's choice between X and Y inferred by LS and by CJ items can be categorized into

- concordant if $L_{iXY} = C_{iXY}$,

Table 1
Notations of Scores or Choices of the LS and CJ Items Concerning Alternatives (X, Y) for an Individual i.

Condition	LS		CJ
	X	Y	(X, Y)
1	L_{iX1}	L_{iY1}	C_{iXY1}
2	L_{iX2}	L_{iY2}	C_{iXY2}
3	L_{iX3}	L_{iY3}	C_{iXY3}
⋮	⋮	⋮	⋮
K	L_{iXK}	L_{iYK}	C_{iXYK}
Scale	L_{iX}	L_{iY}	
Choice	L_{iXY}		C_{iXY}

- discordant if $L_{iXY} \times C_{iXY} = -1$, or
- undetermined if either $L_{iXY} = 0$ or $C_{iXY} = 0$ (but not both).

Collectively, we can assess the degree of congruence between LS and CJ on (X, Y) among all participants by calculating the *proportion* of participants falling into each of the three categories: they are P_C for the concordant, P_D for the discordant, and P_U for the undetermined cases. Note that the sum of $P_D + P_U$ serves as an index measure of effect size (Kelley and Preacher, 2012, p. 140) for our analysis, as it represents the proportion of participants whose inferred choices from LS are *not* congruent with those from CJ.

Whether L_{iXY} and C_{iXY} could be regarded as reliable indicators of LS and CJ is an empirical issue. This can be done by testing whether each derived choice of LS (L_{iXYk}) and each choice of CJ (C_{iXYk}) on the item level is consistent with the aggregated indicator of L_{iXY} and C_{iXY}. In the case of LS, for each participant i and condition k, we can assess the item-total consistency by comparing each pair (L_{iXk}, L_{iYk}) using the discrepancy rule, and counting how many of the K pairs yield the same result as does (L_{iX}, L_{iY}). A similar procedure applies to the CJ case but using the majority rule for assessing the item-total consistency. The above machinery was implemented in our analysis of real data, which we now describe.

1.2. Present study: academic attributions

One major concern in research of achievement motivation is about the individual differences of causal attributions of academic success and failure. There has been a debate on the appropriate methods of assessing academic attributions (Elig and Frieze, 1979; Maruyama, 1982). Three types of methods for measuring attributions were identified by Elig and Frieze: unstructured/open-ended question, structured/independent unipolar rating scales (e.g., LS), and structured/ipsative scales (e.g., percentage assessment and CJ). The percentage assessment (PA) method assumes that the total cause of an event can be parceled out to several specific causes and thus requests participants to individually assign a percentage to each of the possible attributions provided by the researchers. Elig and Frieze (1979) concluded that rating scale was slightly better than PA because the factors revealed by rating scales in exploratory factor analysis (EFA) were more easily interpreted, and thus implied better construct validity. Later, since some major attribution categories had already been presumed and embedded in the items in Elig and Frieze's assessment, Maruyama (1982) reanalyzed Elig and Frieze's data by using confirmatory factor analysis (CFA) on both the assigned percentages and the ratings to the more prominent attribution categories in the assessment. The result of CFA supported the presumed factor structure of attribution categories both for the ratings and for the assigned percentages. Moreover, since the ratings and the assigned percentages were simultaneously included in the same CFA, two "method factors" (i.e., the "scale method factor" and the "percent method factor") also revealed. For the scale method factor, all LS items in different attribution categories, no matter if the item was related to ability, stable effort, unstable effort, task difficulty, luck, or mood, loaded positively in this factor, indicating possible response biases while making LS ratings. By contrast, not all items of PA loaded positively on the percent method factor. The researcher then inferred that the ipsative approach was better than the rating approach for not having the above response biases.

Given the above findings, one question emerges concerning using LS to infer preference. That is, with the problem of response bias while using rating scale to assess attributions, whether the result of LS can accurately reflect individuals' psychological tendency when various attributions are closely competing at the same time?

Take effort and ability as an example. Effort and ability are the two most dominant causal attributions for achievement, even in the responses to open-ended questions (Weiner, 1985, 2018). Attributions of academic success (failure) to ability or effort may induce different emotions (Brown and Weiner, 1984), learning behaviors (e.g., Dweck and Reppucci, 1973), senses of contingent self-worth (Kamins and Dweck, 1999), and academic performance (Mueller and Dweck, 1998). The relative dominance between effort and ability attributions is therefore of both empirical and theoretical interest (e.g., Nicholls, 1976). While using LS to understand the preference of an individual's attribution, suppose this person has attributed both ability and effort as the cause for an academic performance with only slightly different degree. Even though one type of attribution might still be more dominant, with limited number of response categories for LS (say, from 1 to 7) as well as the response bias of that particular individual, the same response category may be assigned to best describe the degree of both effort and ability, resulting in an LS tie. That is, LS may have its limitation in discriminating the subtle difference of favorability between two choices such as ability and effort.

To make the measurement issue concerning effort and ability a bit more complicated, these two types of attributions may function differently or even in reversed directions in success versus failure contexts. For example, according to Weiner and his colleagues' original postulation (Weiner et al., 1972), effort is an internal and unstable (i.e., temporary) while ability is an internal but stable cause of outcome. Therefore, when one has failed an exam, if the poor grade was caused by lack of effort, this individual can always ameliorate the situation by making more effort and gain a better grade in the next exam. Under such circumstances, we can surely say "lack of effort" is a self-serving and thus more popular attribution for most people at least in failure situations (Mezulis et al., 2004). However, to attribute one's own success to lots of hard work is not necessarily self-serving, because it implies that the success is transient and the good grade may not necessarily be replicable. Therefore, while investigating academic attribution, causal attributions for success and failure incidents should be analyzed separately.

Although the aforementioned difference of attribution tendency in success and failure scenarios seemed conceivable, it might not apply universally to people with different cultural backgrounds. Tweed and Lehman (2002) proposed a Confucian-Socratic framework to explain cultural differences in beliefs of learning, the former emphasizes memorization, repetition, and effortful learning while the latter repeated questioning and self-generated knowledge. Specifically, in culturally-Chinese (Confucian) societies, diligence is considered a major and powerful means that can compensate for incompetence and can be used to cultivate the self to proximate intellectuals' ideal state of sagehood (Li, 2003), which is deemed by Confucius to be the moral perfection worth pursuing for all human beings (Lee, 1996). That is, in the landscape of Confucian philosophy, "working hard with unremitting assiduity" is regarded as the core concept of learning as well as a moral imperative. Consequently, not only to avoid attributing academic success to ability saves one from being blamed for lack of modesty but also to attribute success to hardworking demonstrates one's respect for the importance of effort.

In summary, the main goal of this study is to examine whether it is appropriate to use quantitative indicators derived from the results of LS to infer choices among alternatives, with which we used the (binary) CJ procedure (Thurstone, 1927) as a direct measurement of choice decision to compare. We are especially concerned about how participants use LS, which inevitably contains limited number of response categories, to demonstrate their preference when the tendency to favor either one of the two competing choices is near a tie. Our analyses of the concordant, discordant, and undetermined cases will render a demonstration of whether the LS ties imply indecisiveness in terms of academic attribution or the subtle differences of favorability between effort and ability which LS is not capable to reveal.

Since the connotations of making effort attribution in contexts of success and failure are different, just as the meanings of ability attribution after success and failure are not compatible, we conducted the analysis for success and failure scenarios separately. Our participants were students from junior and senior high schools because elementary-school children have not yet developed the full-fledged ability to differentiate the concepts of effort and ability (Nicholls, 1978). Due to the fact that the participants were from a culturally-Chinese society that follows Confucian tradition, we also tested whether they would mainly attribute success to ability as expected from a self-serving rationale or to effort based on Confucian teachings.

2. Method

2.1. Sampling and participants

In order to have a representative sample of junior high and high school students in Taiwan, we used stratified sampling with three strata: junior high schools, "grade A" senior high schools, and "grade B" senior high schools.[1] A total of 998 students, including 332 junior-high-school students and 355 "grade A" and 311 "grade B" high-school students, participated in this study. Only the 929 participants (431 girls and 498 boys; M_{age} = 16.07 years; SD = 1.58) who have filled out all of the 18 items in the inventory were included in the data analysis. We used the upper bound (which occurs when the proportion is 0.5) in the computation of standard errors of estimated proportions, and so the above sample size guarantees that all the 95% confidence intervals of estimated proportions reported in this study are well within ± 0.033.

2.2. Measures: academic attribution inventory

There were 18 items in this inventory. Twelve of them were LS items, and six of them were CJ items. The inventory consisted of two different scenarios, describing a situation that one has performed very well in a test and received an excellent grade and a situation that one has performed very badly and received a terrible grade. After reading each scenario, the participants were then first to fill out three items about the reasons why they think that they had performed so well (so badly). The three items included two seven-point LS items ranging from 1 (strongly disagree) to 7 (strongly agree) about one's attributions of effort and ability followed by one CJ item that required the participant to choose between effort and ability as the major reason for success (failure). Participants were then presented with the same three items but were asked to answer according to how they would respond if were asked by the teacher about the reason of their success (failure). The same three items were then presented again to respond according to how the participants would tell their classmates. Consequently, there were in total nine items (six LS items and three CJ items[2]) following each scenario. The complete list of items is displayed in the Appendix. The protocol of the study followed the ethical principles enunciated by the Taiwanese Psychological Association, and was approved by the Research Ethics Committee in the Department of Psychology, National Taiwan University.

[1] In Taiwan, junior high school students are assigned to different senior high schools based on their scores from a nation-wide high school entrance examination. Therefore, we included both "grade A" and "grade B" senior high school students in the study to reduce sampling bias.

[2] Thus there is no CJ tie in our study.

3. Results

In the following analyses, we first calculated the 95% confidence intervals (CI) for each of the parameters. Next, each of the null hypotheses H_0: $\theta = \theta_0$ was tested by the standard of 0.05 significance level and by whether the 95% CI contains θ_0 (Casella and Berger, 2002).

There are multiple methods for constructing CIs for various estimators. In this study, we used specific statistical procedure for each indicator to estimate its CI (Morey et al., 2016). Specifically, for proportions we computed the Wilson confidence limits (Wilson, 1927) for its CI. For Pearson's correlations we computed the CI based on Fisher's z transformation (Fisher, 1934). Concerning the means, variances, skewness, kurtosis, and McDonald's ωs we simulated the CI for each of them by the bias-corrected and accelerated bootstrap with 10000 bootstrap samples (Efron, 1987).

We mainly used the SAS/STAT 15.1 software (SAS Institute Inc, 2019) to conduct statistical analysis. In particular, we used PROC SURVEYFREQ to estimate the Wilson confidence limits, PROC FMM to estimate the parameters in the beta-binomial model, and PROC CALIS to conduct confirmatory factor analysis. In addition, we used the Boot package (Canty and Ripley, 2019) in R 4.0.1 (R Core Team, 2020) to estimate bootstrap CIs. For the CIs of McDonald's ωs, we used R's MBESS package (Kelley, 2019).

3.1. Psychometric properties of LS scores

As the choice between X and Y in LS is usually inferred from the discrepancy between the total score of X and Y, we first examined whether the psychometric properties of the discrepancy scores of LS are acceptable. The statistics of the effort attribution scores, ability attribution scores, and discrepancy scores for the success and failure scenarios are presented in Table 2. The 95% CIs of skewness for all variables were between ± 2. The 95% CIs of kurtosis for all variables were between ± 3, suggesting that the variables are not distributed awkwardly. We tested the four-factor structure (i.e., ability attribution after success, effort attribution after success, ability attribution after failure, and effort attribution after failure) based on confirmatory factor analysis using MLSB (which stands for normal-theory Maximum Likelihood parameter estimation and the Satorra-Bentler scale corrections) in PROC CALIS. This method also adjusts the estimate of standard errors by the sandwich formula, which is robust to violations of the multivariate normality assumption in confirmatory factor analysis. The four-factor model was acceptable, Satorra-Bentler scaled $X^2(48) = 198.06$, CFI = .96, SRMR = .03, RMSEA = .06 with 95% CI [.05, .07]. For all of the standardized factor loadings, the lower limits of their 95% CIs were larger than .60.

The Pearson correlation coefficient matrix among the variables of LS is presented in Table 3. The numbers in the diagonal of the matrix are replaced by McDonald's ωs to index internal consistency reliability. The McDonald's ω provides a better estimate of internal consistency reliability than Cronbach's α providing that the LS to be analyzed is unidimensional (McNeish, 2018; Zinbarg et al., 2005). The unidimensionality assumption of each LS was supported by the results of the aforementioned confirmatory factor analysis. For all the variable indicators, the values of McDonald's ωs were higher than 0.83, indicating that the discrepancy scores derived from the LSs of effort and ability attributions were not unreliable as Allen and Yen (1979) was concerned about.

3.2. Choices inferred by LS and CJ items

We first note that the mean item-total consistency of CJ items in the success and failure scenarios were .91 and .95, with 95% CIs [.90, .92] and [.94, .96], respectively, whereas the mean item-total consistency of

Table 2
Descriptive statistics [and 95% confidence intervals] of LS variables.

Variable	M	SD	Skew	Kurtosis	Range
Success					
Ability	3.94	1.32	0.00	−0.05	1.00–7.00
	[3.86, 4.03]	[1.27, 1.38]	[−0.11, 0.12]	[−0.22, 0.15]	
Effort	5.48	1.22	−0.98	1.25	1.00–7.00
	[5.40, 5.56]	[1.16, 1.30]	[−1.15, −0.80]	[0.72, 1.85]	
Discrepancy	1.54	1.96	−0.34	1.05	−6.00–6.00
	[1.41, 1.67]	[1.86, 2.08]	[−0.59, −0.13]	[0.55, 1.71]	
Failure					
Ability	3.40	1.51	0.19	−0.48	1.00–7.00
	[3.30, 3.49]	[1.45, 1.57]	[0.09, 0.29]	[−0.62, −0.32]	
Effort	5.87	1.09	−1.17	1.77	1.00–7.00
	[5.80, 5.94]	[1.03, 1.17]	[−1.42, −0.97]	[0.92, 2.92]	
Discrepancy	2.48	2.09	−0.20	−0.06	−6.00–6.00
	[2.35, 2.62]	[2.00, 2.19]	[−0.43, −0.04]	[−0.48, 0.61]	

Note. $N = 929$.

Table 3
Pearson's correlations and reliabilities [and 95% confidence intervals] of LS variables.

Variable	Success			Failure		
	Ability	Effort	Discrepancy	Ability	Effort	Discrepancy
Success						
Ability	.83					
	[.80, .85]					
Effort	−.18	.85				
	[−.24, −.12]	[.82, .87]				
Discrepancy	−.79	.75	.84			
	[−.81, −.76]	[.72, .77]	[.81, .87]			
Failure						
Ability	.12	−.03	−.09	.91		
	[.05, .18]	[−.01, .03]	[−.16, −.03]	[.89, .92]		
Effort	−.01	.36	.23	−.26	.87	
	[−.07, .05]	[.30, .42]	[.17, .29]	[−.32, −.20]	[.84, .90]	
Discrepancy	−.03	.12	.10	−.50	.53	.89
	[−.10, .03]	[.06, .19]	[.04, .16]	[−.55, −.45]	[.49, .58]	[.87, .91]

Note. $N = 929$.

the choices derived from LS items in the success and failure scenarios were .84 and .91, with 95% CIs [.82, .85] and [.89, .92], respectively. These indices indicated that for the ability and effort attributions, scores of individual items with different statements in LS and CJ mostly matched the results from the respective aggregated LS and CJ.

In the success scenario, the results of LS ratings indicated that the proportion of participants who preferred ability was .137 ($n = 128$), 95% CI [.117, .161], the proportion preferring effort was .767 ($n = 713$), 95% CI [.739, .794], and the proportion showing LS tie was .095 ($n = 88$), 95% CI [.078, .115]. The CJ data showed that the proportion of participants who preferred ability was .149 ($n = 138$), 95% CI [.127, .173], and the proportion preferring effort was .851 ($n = 791$), 95% CI [.827, .873]. In the failure scenario, for LS, the proportion preferring ability, preferring effort, and showing LS tie were .074 ($n = 69$, 95% CI [.059, .093]), 0.846 ($n = 786$, 95% CI [.821, .868]), and .080 ($n = 74$, 95% CI [.064, .099]), respectively. For CJ, the proportion preferring ability and preferring effort were 0.081 ($n = 75$, 95% CI [.065, .100]) and .920 ($n = 854$, 95% CI [.900, .935]), respectively. Combining the above findings, it revealed that a total of 626 participants (proportion = .674, 95% CI [.643, .703]) consistently gave effort higher LS ratings and 731 participants (proportion = .787, 95% CI [.759, .812]) picked effort rather than ability as their major attribution using CJ across the two scenarios.

We now turn to comparing choices inferred by LS and by CJ. In the success scenario, the proportion of the participants who favored one attribution according to answers to the CJ items but preferred another attribution based on the ratings in LS was $P_D = .052$ ($n = 48$), 95% CI [.039, .068]. Meanwhile, the proportion of the participants who favored one attribution according to answers to the CJ items but demonstrated equal preference to the two types of attributions based on the ratings in LS was $P_U = .095$ ($n = 88$), 95% CI [.078, .115]. Lumping together, the proportion of the participants demonstrating different patterns of favorable attribution was $P_D + P_U = .146$ ($n = 136$), 95% CI [.125, .171]. In the failure scenario, the respective estimates of proportion were $P_D = .041$ ($n = 38$), 95% CI [.030, .056], $P_U = .080$ ($n = 74$), 95% CI [.064, .099], and $P_D + P_U = .121$ ($n = 112$), 95% CI [.101, .143].

We regard $P_D + P_U$ as reflecting the proportion of people whose relative dominance between ability and effort attributions inferred from LS and from CJ were not congruent. The above results showed that a non-negligible proportion (more than 1/10) of participants in both the success and the failure scenario fell into this category, indicating a substantial discrepancy between choices inferred by LS and by CJ.

3.3. The undetermined cases

We noticed that among those who demonstrated incongruent choices, about 2/3 of them were those who displayed the undetermined pattern between LS and CJ choices. In addition to considering those who demonstrated equivalent LS scores as the ones who just simply could not consistently display their preferred attribution using LS and CJ, other

Table 4
Results of chi-squared goodness-of-fit tests for the undetermined cases.

Model	Number of choosing "effort" among CJ items				χ^2	*df*	*p*
	0	1	2	3			
Success							
Observed # of participants	15	17	20	36			
Expected # of participants							
Binomial model[a]	11.00	33.00	33.00	11.00	71.15	3	<.001
Beta-binomial model[b]	15.20	16.56	20.49	35.74	0.03	1	.868
Failure							
Observed # of participants	7	7	18	42			
Expected # of participants							
Binomial model[a]	9.25	27.75	27.75	9.25	135.44	3	<.001
Beta-binomial model[c]	6.15	9.79	15.83	42.23	1.21	1	.271

Note.
[a] Binomial distribution with $n = 3$ and $p = .5$.
[b] Beta-binomial distribution with $n = 3$, $\hat{\alpha} = 0.93$ and $\hat{\beta} = 0.56$.
[c] Beta-binomial distribution with $n = 3$, $\hat{\alpha} = 1.28$ and $\hat{\beta} = 0.41$.

explanations for the undetermined cases cannot be overlooked. We now examine an alternative explanation.

There were 88 participants (9.5%) in the success scenario and 74 participants (8.0%) in the failure scenario that were categorized as undetermined. Because all of the undetermined cases demonstrated LS ties, a possible explanation for the undetermined choice patterns is that the ties between effort and ability attributions reflect a "genuine state of indecision" (Kendall, 1945). That is, these participants rated both effort and ability with the same LS value; when they responded to CJ items, they were forced to make choices between two alternatives that were viewed by them as indifferent when using LS. In other words, the undetermined patterns resulting from the combined results of LS and CJ do not imply that LS cannot be used to infer choices. Rather, the forced-choice nature of CJ prevents the participants from expressing that they equally prefer ability and effort attributions. Under such circumstances, since each participant must respond to three CJ items, the number of CJ items answered with effort attribution should be *binomially* distributed with $n = 3$ and $p = .5$. This binomial model was soundly rejected by the Pearson chi-square test for both the success ($\chi^2(3) = 71.15$, $p < .0001$) and failure ($\chi^2(3) = 135.44$, $p < .0001$) scenarios.

The above results indicated that participants showing LS ties do not equally distribute their preferences when they respond to the corresponding CJ items. To further explore this issue, we assumed that each participant having an LS tie has a latent probability of preferring effort attribution over ability. Further, we hypothesized that the distribution of these latent probabilities is beta distributed. Accordingly, the number of CJ items answered with effort attribution is the compound distribution of a binomial distribution and a beta distribution, a so-called *beta-binomial distribution* (Williams, 1975), which has a probability mass function of

$$p(k|n,\alpha,\beta) = \binom{n}{k}\frac{B(k+\alpha, n-k+\beta)}{B(\alpha,\beta)}, \tag{3}$$

where $\alpha > 0$ and $\beta > 0$ are parameters, n is the number of CJ items (in our study), and $B(a, b)$ is the beta function. We performed Chi-squared goodness-of-fit tests using the maximum likelihood estimates (Bishop et al., 1975, p. 348). Results showed that the beta-binomial model for both success and failure scenarios fit the data well: $\chi^2(1) = 0.03$, $p = .86$ for the success scenario and $\chi^2(1) = 1.21$, $p = .27$ for the failure scenario.[3] See Table 4 for the observed and expected number of participants for each cell.

Using the estimated parameter values in the beta-binomial distribution ($\hat{\alpha} = 0.93$, $\hat{\beta} = 0.56$ for the success scenario, and $\hat{\alpha} = 1.28$, $\hat{\beta} = 0.41$ for the failure scenario), we can compute the (latent) probabilities of preferring effort attribution over ability for the participants having LS ties, and thus the proportion of participants deviating from the "genuine state of indecision" can be estimated. We found that for the success scenario, about 92% of participants having LS ties revealed $p_i < .45$ or $p_i > .55$, and 52% of them revealed $p_i < .2$ or $p_i > .8$. Meanwhile, for the failure scenario, about 94% of participants having LS ties revealed $p_i < .45$ or $p_i > .55$, and 63% of them revealed $p_i < .2$ or $p_i > .8$. These results suggested that participants having the same LS scores for effort and ability attributions were not necessarily indifferent to the two attributions, as can be seen in their CJ preferences.

Moreover, the presence of the undetermined cases could be a reflection of the setback of LS in terms of response bias as well as limited number of response categories in LS mentioned in Section 1.2. Our data showed that more than half of the participants showing LS ties consistently chose the same attribution in all three CJ items, suggesting that they do have a preference when they are given a chance to freely pick their relatively dominant achievement attribution. Specifically, among the 88 undetermined cases in the success scenario, 51 (58%) of them consistently chose the same attributions in all three conditions, and the numbers were 49 (66%) out of 74 in the failure scenario (see the "0" and "3" columns under 'Number of choosing "effort" among CJ Items' in Table 4). Moreover, out of the 51 (and 49) cases who showed specific preferences, the majority, 71%, $n = 36$ (and 86%, $n = 42$) of them picked effort, instead of ability, as the cause for their success (and failure).

4. Discussion

In this article, we examined whether it is appropriate to use quantitative indicators derived from LS to infer choices, with which we used the CJ procedure to represent a direct measurement of choice decision to compare. We described and implemented an assessment procedure using real data, in which a total of 929 adolescents reported their effort and ability attributions for academic failure and success using both LS and CJ. We found that, although LS was not explicitly designed to infer

[3] We can treat the binomial model as a restricted model of the beta-binomial model (as $\alpha = \beta$ and approaches infinity). And this binomial model was significantly rejected when comparing to the beta-binomial model: $\chi^2(2) = 71.12$, $p < .0001$ for the success scenario and $\chi^2(2) = 134.23$, $p < .0001$ for the failure scenario.

choices, attribution preference derived from LS is generally consistent with that derived from CJ. However, the proportion of people revealing different choices inferred from LS versus CJ exceeded 1/10 for both the success and failure scenarios, suggesting that inferring psychological preference from LS is not without risk.

We also noticed that the majority of participants displaying incongruent decisions of achievement attribution via LS and CJ showed equivalent LS scores between effort and ability attributions (i.e., the undetermined cases), for which a beta-binomial model seems reasonable to grasp the concept. We tested this model and successfully eliminated the possibility that the tied LS scores in effort and ability attributions actually represent participants' true psychological state. Our results suggested that one should be cautious and not simply consider people having LS ties as being in the "genuine state of indecision."

The measure of inferring the choice derived from LS/CJ responses by aggregating the responses to all items was justified by demonstrating the item-total consistency of choices for both LS and CJ. This newly developed index plays a different role from the traditional psychometric indices such as the item-total correlation and internal consistency reliability. In fact, it is possible to have two 3-item LSs, with each having a nearly perfect item-total correlation and internal consistency reliability, but the mean item-total consistency of choices is low. The reason is that item-total correlation and internal consistency reliability are functions of the variances and covariances among LS items and total scores, while the mean item-total consistency of choices has additionally taken into account the mean difference of the two LS scores.

One limitation of this study is that the proportion of people revealing different choices inferred by LS versus CJ might be underestimated due to the design of the Academic Attribution Inventory. Participants always rated their agreement about attributing condition k to ability/effort by LS immediately before they picked their choice in the same condition by CJ. Thus it is possible that some participants may have intentionally responded to LS and CJ congruently, leading to an underestimation of the proportion of incongruent cases. To reduce possible response biases, adding filler items to the inventory and randomizing the order of items may help distract participants from finding the associations of the items between LS and CJ.

We found that the majority of adolescents in our study tended to attribute their academic performance to effort rather than ability no matter whether the data is derived from CJ or LS and whether in the success or failure context. Furthermore, a large proportion of the participants (67.4% in LS; 78.7% in CJ) adhered to the effort attribution in both scenarios of success and failure and did not switch their attribution preference to find a more self-serving excuse dependent on the particular context. This result, on the one hand, violated the common sense that people tend to make more internal, stable, and global attributions (such as "ability") for positive events (such as a success scenario) in order to make themselves look good and, on the other hand, was inconsistent with the results of meta-analyses on self-serving attributional bias which found that the bias is pervasive across age and culture (Mezulis et al., 2004) including Chinese adolescents (Hu et al., 2016). However, the current result from adolescents in a culturally-Chinese society was concordant with Confucian-oriented view, which regards effort as either a moral imperative (Li, 2003) or at least a social obligation (Chang and Lay, 2018) that everyone should abide by at all times.

To gain further understanding of the latent states involved in the discordant and undetermined cases, it is desirable to have a unified model combining both absolute judgments (ratings in LS) and relative judgments (rankings in CJ). The work by Böckenholt (2004), in which a Thurstonian-type model for 2-point LS was introduced, is promising. The idea is to assume that the fluctuations of data between LS and CJ are the results from the discriminal errors induced by the same latent-process. In principle, we can extend the Böckenholt (2004) model to cover multiple-point LS and to test whether it is adequate for explaining the incongruent cases observed in our study.

It is worth mentioning that in the analysis we used the discrepancy rule (Equation (1)) to transform individual LS scores to preferences. This is different from the common practice in the literature of inferring individual preferences based on comparing the group means of LS scores (e.g., Chen et al., 2009; Rentfrow and Gosling, 2003; Stevenson et al., 1990; Wang et al., 2019). It is possible that there are other rules for aggregating individual LS scores and/or CJ choices, and it is unclear how they would affect our results. There is a caveat though. A recent paper by Regenwetter et al. (2019) provided an argument against taking the face values of Likert-type scales for preference or rank ordering. For example, while being asked about the reason behind good performance, an individual giving "ability" an LS score of 6 and "effort" an LS score of 7 might be indifferent between the two attributions if this individual's threshold of discrimination is more than 1. Following the intuition that a person has a strict preference between a pair of alternatives only when their utility difference is beyond this person's threshold (the idea of "semiorder", Luce, 1956), Regenwetter et al. (2019) developed a response-mechanism embedded semiorder framework to model data from rating scales. One direction of future work is to integrate such modeling approach into our study.

The achievement-attribution inventory we described in the present study only contains two alternatives (i.e., ability and effort). Due to this limitation, the two aforementioned models were not able to be thoroughly tested. To further explore the congruence between LS and CJ, future work could expand the inventory to include more than two types of attributions (e.g., effort, ability, luck, and difficulty; Weiner, 1985). Furthermore, in addition to the three LS/CJ conditions in the original survey (i.e., what they would think about themselves, what they would tell their teacher, and what they would tell their classmates), we could add to the inventory a new condition asking the participants to provide their responses about what they would tell their parents if they obtain a good (poor) grade and thus make $C_{iXY} = 0$ possible. The data will allow for more systematic verification of the congruence between the choices inferred from LS and from CJ.

Credit author statement

Che Cheng: Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing-Draft preparation. **Keng-Ling Lay:** Conceptualization, Methodology, Supervision, Writing-Draft preparation. **Yung-Fong Hsu:** Conceptualization, Methodology, Supervision, Writing-Draft preparation. **Yi-Miau Tsai:** Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

Academic Attribution Inventory

Success Scenario: Imagine that you've got a good grade and performed excellently in an exam

- After the exam, how would you usually consider a situation like this from the bottom of your heart?

Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree

1. I performed well because I had worked hard.

   1 2 3 4 5 6 7

2. I performed well because I am good at it.

   1 2 3 4 5 6 7

3. If you could choose only one reason, which of the above is more like the reason for your good performance? Working hard or being good at it? ☐ Being good at it ☐ Working hard

- After the exam, if the teacher asks you why you've got the good grade, what would you say?

Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree

4. I performed well because I had worked hard.

   1 2 3 4 5 6 7

5. I performed well because I am good at it.

   1 2 3 4 5 6 7

6. If you could choose only one reason, which of the above is more like the answer that you would reply back to your teacher for your good performance? Working hard or being good at it? ☐ Being good at it ☐ Working hard

- After the exam, if one of your classmates asks you why you've got the good grade, what would you say?

Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree

7. I performed well because I had worked hard.

   1 2 3 4 5 6 7

8. I performed well because I am good at it.

1 2 3 4 5 6 7

9. If you could choose only one reason, which of the above is more like the answer that you would reply back to your classmate for your good performance? Working hard or being good at it? ☐ Being good at it ☐ Working hard

Failure Scenario: Imagine that you've got a bad grade and performed terribly in an exam

- After the exam, how would you usually consider a situation like this from the bottom of your heart?

Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree

10. I performed poorly because I had not worked hard.

1 2 3 4 5 6 7

11. I performed poorly because I am not good at it.

1 2 3 4 5 6 7

12. If you could choose only one reason, which of the above is more like the reason for your poor performance? Not working hard or not being good at it? ☐ Not working hard ☐ Not being good at it.

- After the exam, if the teacher asks you why you've got the poor grade, what would you say?

Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree

13. I performed poorly because I had not worked hard.

1 2 3 4 5 6 7

14. I performed poorly because I am not good at it.

1 2 3 4 5 6 7

15. If you could choose only one reason, which of the above is more like the answer that you would reply back to your teacher for your poor performance? Not working hard or not being good at it? ☐ Not working hard ☐ Not being good at it

- After the exam, if one of your classmates asks you why you've got the poor grade, what would you say?

Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree

16. I performed poorly because I had not worked hard.

1 2 3 4 5 6 7

17. I performed poorly because I am not good at it.

1 2 3 4 5 6 7

18. If you could choose only one reason, which of the above is more like the answer that you would reply back to your classmate for your poor performance? Not working hard or not being good at it? ☐ Not working hard ☐ Not being good at it.

References

Allen, M., Yen, W., 1979. Introduction to Measurement Theory. Waveland Press.

Ames, C., Archer, J., 1987. Mothers beliefs about the role of ability and effort in school learning. J. Educ. Psychol. 79 (4), 409–414. https://doi.org/10.1037/0022-0663.79.4.409.

Arrow, K.J., 2012. Social Choice and Individual Values, third ed. Yale University Press https://dx.doi.org/10.12987/9780300186987.

Bartlett, C.J., Heermann, E., Rettig, S., 1960. A comparison of six different scaling techniques. J. Soc. Psychol. 51 (4), 343–348. https://doi.org/10.1080/00224545.1960.9922043.

Bishop, Y.M.M., Fienberg, S.E., Holland, P.W., 1975. Discrete Multivariate Analysis: Theory And Practice. MIT Press.

Böckenholt, U., 2004. Comparative judgments as an alternative to ratings: identifying the scale origin. Psychol. Methods 9 (4), 453–465. https://doi.org/10.1037/1082-989x.9.4.453.

Brown, J., Weiner, B., 1984. Affective consequences of ability versus effort ascriptions: controversies, resolutions, and quandaries. J. Educ. Psychol. 76 (1), 146–158. https://doi.org/10.1037/0022-0663.76.1.146.

Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol. Bull. 56 (2), 81–105. https://doi.org/10.1037/h0046016.

Canty, A., Ripley, B., 2019. Boot: Bootstrap Functions [Computer software], Version 1.3-22. https://cran.r-project.org/web/packages/boot/index.html.

Casella, G., Berger, R.L., 2002. Statistical Inference, second ed. Cengage Learning.

Chang, T.-C., Lay, K.-L., 2018. Do Taiwanese adolescents believe in the moral significance of effort and school performance? Chin. J. Psychol. 60 (3), 151–172. https://doi.org/10.6129/CJP.201809_60(3).0001.

Chen, S.W., Wang, H.H., Wei, C.F., Fwu, B.J., Hwang, K.K., 2009. Taiwanese students' self-attributions for two types of achievement goals. J. Soc. Psychol. 149 (2), 179–193. https://doi.org/10.3200/Socp.149.2.179-194.

Dittrich, R., Francis, B., Hatzinger, R., Katzenbeisser, W., 2007. A paired comparison approach for the analysis of sets of Likert-scale responses. Stat. Model. Int. J. 7 (1), 3–28. https://doi.org/10.1177/1471082x0600700102.

Dweck, C.S., 1986. Motivational processes affecting learning. Am. Psychol. 41 (10), 1040–1048. https://doi.org/10.1037/0003-066x.41.10.1040.

Dweck, C.S., Leggett, E.L., 1988. A social cognitive approach to motivation and personality. Psychol. Rev. 95 (2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256.

Dweck, C.S., Reppucci, N.D., 1973. Learned helplessness and reinforcement responsibility in children. J. Pers. Soc. Psychol. 25 (1), 109–116. https://doi.org/10.1037/h0034248.

Efron, B., 1987. Better bootstrap confidence intervals. J. Am. Stat. Assoc. 82 (397), 171–185. https://doi.org/10.1080/01621459.1987.10478410.

Elig, T.W., Frieze, I.H., 1979. Measuring causal attributions for success and failure. J. Pers. Soc. Psychol. 37 (4), 621–634. https://doi.org/10.1037/0022-3514.37.4.621.

Elliott, E.S., Dweck, C.S., 1988. Goals: an approach to motivation and achievement. J. Pers. Soc. Psychol. 54 (1), 5–12. https://doi.org/10.1037/0022-3514.54.1.5.

Fisher, R.A., 1934. Statistical Methods for Research Workers, fifth ed. Oliver & Boyd.

Gonzálvez, C., Sanmartín, R., Vicent, M., Inglés, C.J., Aparicio-Flores, M.P., García-Fernández, J.M., 2018. Academic self-attributions for success and failure in mathematics and school refusal. Psychol. Sch. 55 (4), 366–376. https://doi.org/10.1002/pits.22117.

Harmon-Jones, C., Summerell, E., Bastian, B., 2018. Anger increases preference for painful activities. Motivation Science 4 (4), 301–314. https://doi.org/10.1037/mot0000090.

Hu, T., Zhang, D., Ran, G., 2016. Self-serving attributional bias among Chinese adolescents. Pers. Indiv. Differ. 91, 80–83. https://doi.org/10.1016/j.paid.2015.10.008.

Joubert, T., Inceoglu, I., Bartram, D., Dowdeswell, K., Lin, Y., 2015. A comparison of the psychometric properties of the forced choice and Likert scale versions of a personality instrument. Int. J. Sel. Assess. 23 (1), 92–97. https://doi.org/10.1111/ijsa.12098.

Kamins, M.L., Dweck, C.S., 1999. Person versus process praise and criticism:Implications for contingent self-worth and coping. Dev. Psychol. 35 (3), 835–847. https://doi.org/10.1037/0012-1649.35.3.835.

Kelley, K., 2019. MBESS: the MBESS R Package (Version 4.7.0) [Computer software]. https://cran.r-project.org/web/packages/MBESS/index.html.

Kelley, K., Preacher, K.J., 2012. On effect size. Psychol. Methods 17 (2), 137–152. https://doi.org/10.1037/a0028086.

Kendall, M.G., 1945. The treatment of ties in ranking problems. Biometrika 33 (3), 239–251. https://doi.org/10.2307/2332303.

Kendall, M.G., 1990. Rank Correlation Methods, fifth ed. E. Arnold.

Lay, K.-L., Tsai, Y.-M., 2005. [Unpublished Raw Data on Moral, Pragmatic, and Entity Concepts of Academic Effort Among Chinese Adolescents: A Developmental Perspective]. National Taiwan University.

Lee, W.O., 1996. The culture context for Chinese learners: conceptions of learning in the Confucian tradition. In: Watkins, D., Biggs, J.B. (Eds.), The Chinese Learner: Cultural, Psychological, and Contextual Influences. Comparative Education Research Centre, pp. 25–41.

Li, J., 2003. U.S. and Chinese cultural beliefs about learning. J. Educ. Psychol. 95 (2), 258–267. https://doi.org/10.1037/0022-0663.95.2.258.

Likert, R., 1932. A technique for the measurement of attitudes. Arch. Psychol. 22 (140), 5–55.

Lord, F., Novick, M., Birnbaum, A., 1968. Statistical Theories of Mental Test Scores. Addison-Wesley.

Luce, R.D., 1956. Semiorders and a theory of utility discrimination. Econometrica 24 (2), 178–191. https://doi.org/10.2307/1905751.

Marsh, H.W., 1984. Relations among dimensions of self-attribution, dimensions of self-concept, and academic achievements. J. Educ. Psychol. 76 (6), 1291–1308. https://doi.org/10.1037/0022-0663.76.6.1291.

Maruyama, G.M., 1982. How should attributions be measured? a reanalysis of data from Elig and Frieze. Am. Educ. Res. J. 19 (4), 552–558. https://doi.org/10.2307/1162543.

McNeish, D., 2018. Thanks coefficient alpha, we'll take it from here. Psychol. Methods 23 (1), 412–433. https://doi.org/10.1037/met0000144.

Mezulis, A.H., Abramson, L.Y., Hyde, J.S., Hankin, B.L., 2004. Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. Psychol. Bull. 130 (5), 711–747. https://doi.org/10.1037/0033-2909.130.5.711.

Morey, R.D., Hoekstra, R., Rouder, J.N., Lee, M.D., Wagenmakers, E.-J., 2016. The fallacy of placing confidence in confidence intervals. Psychonomic Bull. Rev. 23 (1), 103–123. https://doi.org/10.3758/s13423-015-0947-8.

Mueller, C.M., Dweck, C.S., 1998. Praise for intelligence can undermine children's motivation and performance. J. Pers. Soc. Psychol. 75 (1), 33–52. https://doi.org/10.1037//0022-3514.75.1.33.

Nicholls, J.G., 1976. Effort is virtuous, but its better to have ability: evaluative responses to perceptions of effort and ability. J. Res. Pers. 10 (3), 306–315. https://doi.org/10.1016/0092-6566(76)90020-9.

Nicholls, J.G., 1978. The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. Child Dev. 49 (3), 800–814. https://doi.org/10.2307/1128250.

R Core Team, 2020. R: A Language and Environment for Statistical Computing [Computer software], Version 4.0.1. http://www.R-project.org/.

Regenwetter, M., Hsu, Y.-F., Kuklinski, J.H., 2019. Towards meaningful inferences from attitudinal thermometer ratings. Decision 6 (4), 381–399. https://doi.org/10.1037/dec0000106.

Rentfrow, P.J., Gosling, S.D., 2003. The do re mi's of everyday life: the structure and personality correlates of music preferences. J. Pers. Soc. Psychol. 84 (6), 1236–1256. https://doi.org/10.1037/0022-3514.84.6.1236.

SAS Institute Inc, 2019. SAS/STAT software (Version 15.1) [Computer software]. http://www.sas.com/.

Stevenson, H.W., Lee, S.-Y., Chen, C., Stigler, J.W., Hsu, C.-C., Kitamura, S., Hatano, G., 1990. Contexts of achievement: a study of American, Chinese, and Japanese children. Monogr. Soc. Res. Child Dev. 55 (1–2), 1–123. https://doi.org/10.2307/1166090.

Thurstone, L.L., 1927. Psychophysical analysis. Am. J. Psychol. 38 (3), 368–389. https://doi.org/10.2307/1415006.

Tsujimoto, K.C., Frijters, J.C., Boada, R., Gottwald, S., Hill, D., Jacobson, L.A., Lovett, M. W., Mahone, E.M., Willcutt, E.G., Wolf, M., Bosson-Heenan, J., Gruen, J.R., 2018. Achievement attributions are associated with specific rather than general learning delays. Learn. Indiv Differ 64, 8–21. https://doi.org/10.1016/j.lindif.2018.04.002.

Tweed, R.G., Lehman, D.R., 2002. Learning considered within a cultural context Confucian and Socratic approaches. Am. Psychol. 57 (2), 89–99. https://doi.org/10.1037/0003-066x.57.2.89.

Vuletich, H.A., Kurtz-Costes, B., Bollen, K.A., Rowley, S.J., 2018. A longitudinal study of the domain-generality of African American students' causal attributions for academic success. J. Educ. Psychol. 111 (4), 459–474. https://doi.org/10.1037/edu0000299.

Wang, X., Smith, C., Ashley, L., Hyland, M.E., 2019. Tailoring self-help mindfulness and relaxation techniques for stroke survivors: examining preferences, feasibility and acceptability. Front. Psychol. 10 (391) https://doi.org/10.3389/fpsyg.2019.00391.

Weiner, B., Kukla, A., 1970. An attributional analysis of achievement motivation. J. Pers. Soc. Psychol. 15 (1), 1–20. https://doi.org/10.1037/h0029211.

Weiner, B., 1985. An attributional theory of achievement-motivation and emotion. Psychol. Rev. 92 (4), 548–573. https://doi.org/10.1037/0033-295x.92.4.548.

Weiner, B., 2018. The legacy of an attribution approach to motivation and emotion: a no-crisis zone. Motivation Science 4 (1), 4–14. https://doi.org/10.1037/mot0000082.

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S., Rosenbaum, R.M., 1972. Perceiving the causes of success and failure. In: Jones, E.E., Kanouse, D.E., Kelley, H.H., Nisbett, R. E., Valins, S., Weiner, B. (Eds.), Attribution: Perceiving the Causes of Behavior, first ed. General Learning Corporation, pp. 95–120.

Williams, D.A., 1975. The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. Biometrics 31 (4), 949–952. https://doi.org/10.2307/2529820.

Wilson, E.B., 1927. Probable inference, the law of succession, and statistical inference. J. Am. Stat. Assoc. 22 (158), 209–212. https://doi.org/10.2307/2276774.

Zinbarg, R.E., Revelle, W., Yovel, I., Li, W., 2005. Cronbach's α, Revelle's β, and McDonald's ω_h: their relations with each other and two alternative conceptualizations of reliability. Psychometrika 70 (1), 1–11. https://doi.org/10.1007/s11336-003-0974-7.