Research Article

An Analysis of High and Low Intercorrelations between Mathematics Self-Efficacy, Anxiety, and Achievement Variables: A Prerequisite for a Reliable Factor Analysis

Erik Bergqvist (ID)**, Timo Tossavainen, and Maria Johansson**

Lulea University of Technology, Faculty of Humanities and Social Sciences, Lulea, Sweden

Correspondence should be addressed to Erik Bergqvist; erik.bergqvist@ltu.se

This paper draws on data from a quantitative study of upper secondary students' general mathematical self-efficacy, anxiety towards mathematics, and their relationship to achievement in mathematics. The main objective of this article is to discuss the type of information that may be lost if potential problems of validity and extreme multicollinearity in exploratory factor analysis would be solved by only removing variables without doing a profound analysis. We also describe a method that treats Likert items in the questionnaire as ordinal variables that may represent the underlying continuous variable. Our study shows, for example, that removal of problematic variables without a profound analysis leads to a loss of significant information about test anxiety. Our qualitative analysis of problematic variables also led to an unexpected finding regarding the relationship between general mathematical self-efficacy and motivational values in mathematics.

1. Introduction

In social cognitive theory, self-efficacy beliefs play an important role in students' learning, and mathematics education research has many times shown a positive correlation between perceived mathematics self-efficacy and mathematics achievement (e.g., [1–3]).

Most quantitative studies of students' mathematical self-beliefs are analyzed using exploratory factor analysis (EFA) or similar methods, which often are based on correlations. Despite this fact, it is difficult to find research articles that include a discussion on the "factorability" of the correlations. For example, extreme multicollinearity is considered problematic and particularly common in data coming from social science fields such as education and didactics. Yet studies often lack an analysis of the severity of multicollinearity and possible consequences of removing variables. Sufficient intercorrelation between variables is essential to conduct EFA. But too highly correlated variables (extreme multicollinearity) are a problem in EFA, since it makes it impossible to determine the unique contribution of

each such variable to a factor [4]. If data show indications of extreme multicollinearity, one approach is to use a quantitative criterion such as $r > 0.80$ and remove one or both of the highly intercorrelating variables. Another issue relating to the "factorability" of the correlations is variables with lots of intercorrelations below 0.30. The removal of these types of variables is equally important because they will not be able to cluster in any of the factors [4, 5].

Research articles on studies using EFA, which also discuss data preparation for EFA, are rare, especially in the research areas relating to self-efficacy beliefs. However, some research articles highlight the value of dealing with extreme multicollinearity. For example, Marsh [6] present and discuss a previously conducted study [7] that had examined the relationships between self-efficacy, self-concept, and achievement in mathematics. In the mentioned study, it was incorrectly claimed that self-efficacy could predict performance best. Marsh [6] explain why the interpretation is incorrect by pointing out that self-efficacy and self-concept were very strongly correlated, the standard errors were large, and the confidence intervals were wide. Thus, there were no

good reasons to claim that neither self-concept nor self-efficacy predicts the achievement best. They concluded that researchers should be very cautious with interpretations of results when extreme multicollinearity is present in data.

This paper draws data from a pilot study aiming at investigating general mathematics self-efficacy and anxiety towards mathematics among upper secondary students in a municipal school in Sweden. Data were collected using a questionnaire consisting of a set of Likert scales measuring the abovementioned latent variables. The participants' responses ($N = 79$) are analyzed using, among other things, EFA.

With this paper, we want to describe and discuss what kind of challenges are related to the preparation of data —especially data of a size that is typical for many doctoral studies in didactics—for a successful EFA, including, among other things, the confronting of possible problems with multicollinearity and construct validity. More specifically, we want to investigate the advantages of including an in-depth analysis of identified problematic variables in EFA, instead of only using a quantitative criterion. Our research questions for this article are as follows:

(1) What kind of information regarding students' mathematics self-beliefs is obtained with the help of an in-depth analysis of variables with lots of low intercorrelation coefficients?

(2) What kind of information regarding students' mathematics self-beliefs is obtained with the help of an in-depth analysis of variables with high intercorrelations?

To answer these questions properly, we first begin to recall the most important prerequisites for using EFA successfully and describe shortly a theoretical framework to study self-beliefs to show what kind of challenges a researcher typically meets within such a framework. Thereafter, we represent our method for the pilot study in more detail. However, since this paper is not a report of the final findings from the pilot study, we only briefly discuss our final factor model. This discussion and the final factor model are contained in a section after the results of the present study.

1.1. Prerequisites for EFA. In the literature, there are many good introductory texts about how to conduct EFA successfully. Examples that fit also small data include [8–10]. Additionally, when issues related to the "factorability" of the correlation matrix, such as treatment of extreme multicollinearity and redundancy, some recommended procedures can be found in, e.g., [5] and in multivariate textbooks such as [11]. However, we found no complete guidelines on how to conduct the pretreatment of multicollinear data, possibly because of many subjective criteria. For example, one heuristic is to look for correlations greater than 0.80 and remove one or several of the highly correlating variables [4, 5]. However, a problem with that kind of heuristics is that two correlations of 0.60 may have the same effect as one of 0.80 [12]. Additionally, the remaining explanatory variables

must always be theoretically grounded, which makes the decision of removing interdependent variables especially important and worth examining thoroughly.

The determinant of the correlation matrix assesses multicollinearity in data. If the determinant is less than the heuristic 0.00001, extreme multicollinearity is evident in the data, and therefore, one of the highly correlating variables should be removed. Further, Bartlett's test can be used to test if the overall correlations are too small. But, even if the test is statistically significant, it is still essential to identify and remove variables that have lots of intercorrelations below 0.30. The motivation for the removal of those variables is that they may not be enough focused to contribute to underlying factors [4, 5].

The basis for EFA or similar methods is correlations. One of the most commonly used correlation methods is the Pearson correlation, which also happens to be the one most incorrectly used in social and behavioral research [13]. Pearson correlations assume that the variables (often Likert items) are continuous (and metric) variables. However, this preassumption is debatable, since ordinal variables are by their nature preferably ordered categorical variables, which makes the metric assumption quite controversial.

One way of dealing with the preassumption of metric variables is using the polychoric correlation, which is the estimation between two normally distributed continuous underlying variables with the aid of two observed ordinal variables [14]. The calculations of polychoric correlations are based on frequency tables, which in some cases also include zero frequency cells. This problem can, however, dealt with by applying a continuity correction, i.e., the zeros replaced by a small number such as 0.5. But the presence of many such corrections leads to unstable correlations, and this problem is particularly common in data with small sample sizes [15]. Therefore, continuity correction should be used with caution in data with small sample sizes. The correlation between Likert objects with more than six response categories has shown almost the same correlation coefficients whether using Pearson or polychoric correlation [16]. Thus, the use of Pearson correlations with ordinary data can be discussed (c.f. [17]). Still, the methodology issue of EFA with ordinal data based on Pearson correlations remains [13].

Concerning studies with small sample sizes, for example, in the field of education and didactics, there is a continuous debate on the necessary sample size to conduct an adequate EFA. A common "rule of thumb" is that a minimum of 10 participants per variable is required [4, 5]. However, EFA could still be possible and reasonable with small sample sizes, because the evaluation of reliable factor model, i.e., if the factor model is a good recovery of the population solution, mainly depends on the communalities of the variables and the number of variables per factor [18–22].

1.2. Theoretical Framework. Perceived self-efficacy is defined as a personal judgment of "how well one thinks one can execute courses of action required to deal with prospective situations" [23]. Further, mathematics self-efficacy is defined as an individual's belief about his/her ability in mathematics,

i.e., the strength of the confidence in one's ability to accomplish a mathematical task. Self-efficacy shares both similarities and differences with another psychological construct, self-concept. Self-concept is a perception of oneself, a self-description of the physical, psychological, and social self. When self-efficacy focuses on an individual's ability in completing a specific task and concerns future-oriented conceptions, self-concept has a stronger relation to the social environment and involves the judgment of the self that is past-oriented [24, 25]. Both constructs are an important part of the overall perception of self [2].

2. Method

The data collection for this study was conducted at a municipal high school in Sweden, which has a total of about 400 students. The self-reporting questionnaire was administered to students online during four regular math lessons. The questionnaire included altogether 29 five-step Likert items assessing a respondent's general mathematics self-efficacy and anxiety beliefs. The participating students are sixteen or seventeen years old, and they are studying at a university entrance qualification program. All participants participated in the study voluntarily. Seventy-nine students, including first- and second-year students, chose to participate in the study. The number of girls was 57, and the number of boys was 22.

2.1. Measures.

The underlying constructs of students' general mathematical self-efficacy and anxiety beliefs were assessed using a Swedish adaptation of a previous questionnaire, the Mathematical Self-Efficacy and Anxiety Questionnaire (MSEAQ).

The MSEAQ has been developed by May [26]. The pilot version of the original questionnaire consisted of items based on research literature and adapted items from previous questionnaires designed to measure mathematics self-efficacy and anxiety towards mathematics. The items that ended up in the final version were designed acknowledging the feedback received at a conference and from a group of college students in a series of interviews conducted during the pilot. Finally, May [26] conducted an EFA based on data from 109 college students. Thirteen students were interviewed to help with the interpretation of the retained factors. This led, for instance, to that the item "I believe I can think like a mathematician" was removed from the data before the final EFA, since students' interviews revealed that the students had not interpreted this item consistently. In this study, however, we chose to include that item in the questionnaire. The other 28 items in the original MSEAQ include thirteen self-efficacy scales (MSEAQ-SE) and fifteen anxiety scales (MSEAQ-A). The original MSEAQ was considered highly reliable in terms of internal consistency: for MSEAQ-SE, $\alpha = 0.90$, and for MSEAQ-A, $\alpha = 0.91$. These reliability estimates are based on the responses from 61 students.

According to May [26], the conducted EFA resulted in a five-factor model: mathematics self-efficacy, grade anxiety, future, in-class anxiety, and assignment factor. The first factor represents a general mathematics self-efficacy (e.g., "I believe I can get an "A" when I am in a mathematics course"), factor 2 relates to anxieties toward grades in mathematics (e.g., "I worry that I will not be able to do well on mathematics tests"), factor 3 relates to anxieties regarding future career and mathematics courses (e.g., "I get nervous when I have to use mathematics outside of school"), factor 4 represents in-class anxiety (e.g., "I am afraid to give an incorrect answer during my mathematics class"), and finally, factor 5 represents abilities and anxieties to complete assignments (e.g., "I believe I can complete all of the assignments in a mathematics course").

Table 1 presents the translation of the MSEAQ with the original wordings in italics. Each statement was followed by a five-point Likert scale (English wordings in italics): 1 = aldrig (*never*), 2 = sällan (*seldom*), 3 = ibland (*sometimes*), 4 = ofta (*often*), and 5 = vanligtvis (*usually*). In the original MEAQ, the "no response" option was also available. However, in our study, that option was not included because students were assumed to be able to assess statements concerning his/her mathematics self-beliefs.

When a questionnaire is translated into another language, there is always a risk of decreased reliability [27] which may even raise a need to change the factor model. This risk can, however, be analyzed and the possible problems can be solved by using confirmatory factor analysis (CFA). In the present study, a relevant question is whether the five-factor model of May [26] fits our data. Thus, a CFA was conducted using the lavaan package version 0.6-7 [28] in R version 4.0.2 [29]. In general, Likert items are considered ordered categorical. Consequently, diagonally weighted least squares (DWLS) with polychoric correlations and Satorra–Bentler correction were implemented to estimate the model parameters. The five-factor model fit was reasonably well (see [30] for cutoff values), with a TLI of 0.989 (robust TLI 0.978), although the RMSEA index was almost within an acceptable range with 0.079 (robust 0.114), 90% CI (0.064, 0.094), and robust CI (0.106, 0.122). If we aimed to report the final factor model, the factor model of May [26] would hence need some improvements to be classified to have a good fit. However, for this study, the above indices are sufficient to show that the items used in this study are adequate. Besides, the translated mathematics self-efficacy subscale (se) and mathematics anxiety subscale (an) were considered highly reliable in terms of internal consistency: for self-efficacy, $\alpha = 0.92$, and for anxiety, $\alpha = 0.91$.

2.2. Data Analysis.

The sample consisted of 9% missing values and was assumed to be missing completely at random. One often-used strategy is to remove missing values completely (listwise/pairwise), if a maximum of 10% of the values is missing. However, that kind of missing data strategy can often lead to a loss in statistical power and biased standard errors, especially if removing a large part of the sample. Another frequently used method is single imputation (mean or regression), but this method often results in an underestimation of variances and overestimation of correlations

TABLE 1: The original items of the MSEAQ and their Swedish translations.

Variable	Item
se1	Jag känner mig tillräckligt självsäker för att ställa frågor under matematiklektionen *I feel confident enough to ask questions in my mathematics class*
an2	Jag blir spänd när jag förbereder mig inför ett matematikprov *I get tense when I prepare for a mathematics test*
an3	Jag blir nervös när jag är tvungen att använda mina matematikkunskaper utanför skolan *I get nervous when I have to use mathematics outside of school*
se4	Jag tror det ska gå bra på matematikprov *I believe I can do well on a mathematics test*
an5	Jag oroar mig för att det inte kommer att gå bra när jag behöver använda mina matematikkunskaper i mitt framtida yrke *I worry that I will not be able to use mathematics in my future career when needed*
an6	Jag oroar mig för att jag inte kommer att få ett bra betyg i matematikkursen *I worry that I will not be able to get a good grade in my mathematics course*
se7	Jag tror jag kommer att bli klar med alla uppgifter i matematikkursen *I believe I can complete all of the assignments in a mathematics course*
an8	Jag oroar mig för att det inte kommer att gå bra på matematikproven *I worry that I will not be able to do well on mathematics tests*
se9	Jag tror jag är en person som är bra på matematik *I believe I am the kind of person who is good at mathematics*
se10	Jag tror det kommer att gå bra när jag behöver använda mina matematikkunskaper i mitt framtida yrke *I believe I will be able to use mathematics in my future career when needed*
an11	Jag känner mig stressad av att lyssna på min matematiklärare *I feel stressed when listening to mathematics instructors in class*
se12	Jag tror jag kan förstå matematikkursens innehåll *I believe I can understand the content in a mathematics course*
se13	Jag tror jag kan få betyget A i matematikkursen *I believe I can get an "A" when I am in a mathematics course*
an14	Jag blir nervös när det ställs frågor under matematiklektionen *I get nervous when asking questions in class*
an15	Jag blir stressad av att jobba med matematikuppgifter hemma *Working on mathematics homework is stressful for me*
se16	Jag tror det ska gå bra att lära mig matematik *I believe I can learn well in a mathematics course*
an17	Jag oroar mig för att jag inte kommer att kunna tillräckligt mycket för att göra bra ifrån mig i nästa matematikkurs *I worry that I do not know enough mathematics to do well in future mathematics courses*
an18	Jag oroar mig för att jag inte kommer att bli klar med alla uppgifter i matematikkursen *I worry that I will not be able to complete every assignment in a mathematics course*
se19	Jag känner mig självsäker när jag gör matematikprov *I feel confident when taking a mathematics test*
se20	Jag tror jag är en person som kan använda matematik *I believe I am the type of person who can do mathematics*
se21	Jag känner att det kommer att gå bra i nästa matematikkurs *I feel that I will be able to do well in future mathematics courses*
an22	Jag oroar mig för att jag inte kommer att förstå matematiken som finns med i kursen *I worry I will not be able to understand the mathematics*
se23	Jag tror jag kan använda matematiken i kursen *I believe I can do the mathematics in a mathematics cours*
an24	Jag oroar mig för att jag inte kommer att få betyget A i matematikkursen *I worry that I will not be able to get an "A" in my mathematics course*
an25	Jag oroar mig för att jag inte kommer att lära mig bra i matematikkursen *I worry that I will not be able to learn well in my mathematics course*
an26	Jag blir nervös när jag gör matematikprov *I get nervous when taking a mathematics test*
an27	Jag är rädd för att säga fel under matematiklektionen *I am afraid to give an incorrect answer during my mathematics class*
se28	Jag tror jag kan tänka som en matematiker *I believe I can think like a mathematician*
se29	Jag är självsäker när jag använder mina matematikkunskaper utanför skolan *I feel confident when using mathematics outside of school*

Note: mathematics self-efficacy (se) and mathematics anxiety (an) items.

TABLE 2: The correlation matrix.

	an1	an2	an3	se4	an5	an6	se7	an8	se9	se10	an11	se12	se13	an14	an15	se16	an17	an18	se19	se20	se21	an22	se23	an24	an25	an26	an27	se28	se29
an1		-0.18	-0.26*	0.44‡	-0.37‡	-0.32‡	0.33†	-0.28*	0.30†	0.35‡	-0.22	0.34†	0.13	-0.59‡	-0.15	0.23*	-0.31†	-0.17	0.31†	0.38‡	0.28*	-0.25*	0.27*	-0.15	-0.34†	-0.27*	-0.52‡	0.07	0.44‡
an2			0.44‡	-0.40‡	0.56‡	0.46‡	-0.04	0.75‡	-0.56‡	-0.23	0.30†	-0.12	-0.44‡	0.53‡	0.49‡	-0.37‡	0.45‡	0.29†	-0.59‡	-0.55‡	-0.47‡	0.49‡	-0.23	0.34†	0.41‡	0.85‡	0.46‡	-0.70‡	-0.52‡
an3				-0.26*	0.53‡	0.38‡	-0.17	0.45‡	-0.48‡	-0.42‡	0.20	-0.22	-0.25*	0.41‡	0.19	-0.31†	0.50‡	0.20	-0.48‡	-0.40‡	-0.43‡	0.36‡	-0.29‡	0.09	0.33†	0.42‡	0.54‡	-0.48‡	-0.57‡
se4					-0.49‡	-0.58‡	0.52‡	-0.63‡	0.64‡	0.52‡	-0.30†	0.60‡	0.66‡	-0.39‡	-0.38‡	0.72‡	-0.59‡	-0.51‡	0.72‡	0.60‡	0.70‡	-0.63‡	0.49‡	-0.12	-0.48‡	-0.48‡	-0.29†	0.58‡	0.59‡
an5						0.58‡	-0.41‡	0.61‡	-0.53‡	-0.68‡	0.31†	-0.38‡	-0.28*	0.42‡	0.46‡	-0.32‡	0.62‡	0.48‡	-0.59‡	-0.58‡	-0.60‡	0.56‡	-0.25*	0.29*	0.50‡	0.69‡	0.41‡	-0.46‡	-0.62‡
an6							-0.48‡	0.69‡	-0.51‡	-0.59‡	0.44‡	-0.53‡	-0.39‡	0.39‡	0.39‡	-0.64‡	0.63‡	0.66‡	-0.62‡	-0.52‡	-0.77‡	0.66‡	-0.54‡	0.46‡	0.73‡	0.51‡	0.46‡	-0.45‡	-0.52‡
se7								-0.30†	0.47‡	0.60‡	-0.36†	0.70‡	0.37‡	-0.22	-0.37‡	0.43‡	-0.55‡	-0.70‡	0.39‡	0.52‡	0.45‡	-0.45‡	0.35‡	-0.02	-0.27*	-0.13	-0.22	0.28*	0.43‡
an8									-0.67‡	0.42‡	0.42‡	-0.37‡	-0.53‡	0.60‡	0.55‡	-0.68‡	0.67‡	0.61‡	-0.80‡	-0.63‡	-0.74‡	0.72‡	-0.49‡	0.36†	0.55‡	0.78‡	0.49‡	-0.62‡	-0.68‡
se9										0.58‡	-0.34‡	0.45‡	0.57‡	-0.55‡	-0.52‡	0.64‡	-0.55‡	-0.61‡	0.67‡	0.62‡	0.68‡	-0.63‡	0.49‡	-0.12	-0.52‡	-0.55‡	-0.38‡	0.75‡	0.60‡
se10											-0.24*	0.63‡	0.39‡	-0.30†	-0.39‡	0.51‡	-0.67‡	-0.57‡	0.55‡	0.66‡	0.57‡	-0.56‡	0.44‡	-0.21	-0.51‡	-0.35‡	-0.29†	0.51‡	0.61‡
an11												-0.39‡	-0.10	0.44‡	0.42‡	-0.33‡	0.46‡	0.46‡	-0.34‡	-0.43‡	-0.39‡	0.58‡	-0.36‡	0.25*	0.44‡	0.33†	0.34†	-0.21	-0.43‡
se12													0.32†	-0.23	-0.33†	0.50‡	-0.62‡	-0.61‡	0.46‡	0.59‡	0.52‡	-0.66‡	0.55‡	-0.21	-0.40†	-0.17	-0.15	0.31*	0.35‡
se13														-0.28*	-0.37‡	0.45‡	-0.56‡	-0.48‡	0.72‡	0.44‡	0.62‡	-0.50‡	0.50‡	0.16	-0.27*	-0.39‡	-0.28*	0.58‡	0.39‡
an14															0.13	-0.31‡	0.36†	0.32†	-0.45‡	-0.53‡	-0.50‡	0.60‡	-0.31†	0.29*	0.44‡	0.48‡	0.70‡	-0.45‡	-0.62‡
an15																-0.43‡	0.54‡	0.63‡	-0.54‡	-0.47‡	-0.56‡	0.38‡	-0.24*	0.18	0.49‡	0.52‡	0.00	-0.41‡	-0.43‡
se16																	-0.56‡	-0.50‡	0.54‡	0.53‡	0.70‡	0.60‡	0.61‡	-0.24*	-0.55‡	-0.34‡	-0.20	0.57‡	0.46‡
an17																		0.68‡	-0.70‡	-0.68‡	-0.72‡	-0.68‡	-0.44‡	0.34†	0.55‡	0.54‡	0.35‡	-0.57‡	-0.61‡
an18																			-0.51‡	-0.65‡	-0.67‡	0.79‡	-0.43‡	0.24*	0.53‡	0.34‡	0.18	-0.47‡	-0.53‡
se19																				0.63‡	0.76‡	0.71‡	0.39‡	-0.17	-0.55‡	-0.67‡	-0.41‡	0.61‡	0.61‡
se20																					0.65‡	-0.68‡	0.42‡	-0.29*	-0.52‡	-0.57‡	-0.37‡	0.64‡	0.69‡
se21																						-0.82‡	0.60‡	-0.27*	-0.76‡	-0.57‡	-0.31†	0.61‡	0.58‡
an22																							-0.54‡	0.36†	0.63‡	0.56‡	0.32†	-0.57‡	-0.48‡
se23																								-0.06	-0.45‡	-0.23	-0.22	0.39‡	0.44‡
an24																									0.45‡	0.35‡	0.24*	-0.30†	-0.37‡
an25																										0.53‡	0.41‡	-0.45‡	-0.48‡
an26																											0.55‡	-0.58‡	-0.63‡
an27																												-0.22	-0.46‡
se28																													0.62‡

Note. N = 79. Significance levels ‡ p < 0.001, † p < 0.01, and * p < 0.05 (two-tailed).

[31]. A recommended strategy for dealing with missing data is, however, multiple imputations, and that was applied to these data using mice version 3.8.0 [32] with a multinomial logit model in R version 4.0.2 [29].

All statistical calculations were performed using the statistical software R version 4.0.2 [29]. Variables with high and low polychoric intercorrelations were selected. More precisely, polychoric correlations were calculated using the psych package version 2.0.8 [33] in R version 4.0.2 [29], and the correlation matrix was scanned for intercorrelations $|r| > 0.60$ and $|r| < 0.30$. These items were analyzed to identify which of these items having validity and multicollinearity issues. Further, a qualitative analysis was conducted to explore what kind of information that might be lost if the identified problematic items were removed from the data. More precisely, the qualitative analyses focused on investigating the content of each item and comparing which aspects or relations between the latent variables that are covered or lost depending on which items that would be included or excluded from the factor model.

3. Results and Discussion

In the following sections, we present the results by first discussing the low and high intercorrelations between variables and answering the research questions, and then describing the final factor model.

3.1. Low Intercorrelations between Variables. Table 2 presents the correlation matrix, and it shows that variable an24 has lots of low $|r| < 0.30$ intercorrelations between variables. However, variable an24 has a few moderate statistically significant intercorrelations, e.g., the correlation with variable an6 ($r = 0.46$), variable an8 ($r = 0.36$), and variable an25 ($r = 0.45$). When students assessed statement an24, more than 50% of the students reported that they never or seldom worry that they will not be able to get an "A" in the mathematics course. In contrast to an24, more than 50% of the students reported that they often or usually worry about getting a good grade, or worrying doing well on tests. Additionally, concerning variable an25, more than 50% of the students reported that they sometimes or more frequently think about the statement "I worry that I will not be able to learn well in my mathematics course." This result shows that variable an24 is interpreted differently depending on a student's ambition when it comes to the course grade. However, this may also be an indicator of students' general self-belief in mathematics, since students worry about getting a good grade but not as much to get an "A," the highest grade. Consequently, the removal of variable an24 from EFA without an in-depth analysis would result in a significant loss of information associated with student's mathematics self-beliefs.

Table 2 also contains other low or nonexisting correlations worth mentioning. For example, one might expect that the correlation between variable an24 and variable se13, "I believe I can get an "A" when I am in a mathematics course," ($r = 0.16$, $p > 0.05$) would be negative. Variable an24 is related to mathematics anxiety, and variable se13 is related to mathematics self-efficacy, and previous research has shown that the constructs are strongly correlated. However, the correlation is not statistically significant in these data. This can be considered as a significant finding which however would be lost if variable an24 were automatically removed only by the criteria based on $|r| < 0.30$.

However, the variable an24 is somewhat problematic. Because the interpretation of the statement can take place in many different ways; some students do not pay much attention to their grades but are only interested in passing the course. Other students think that they cannot get the highest grade even if they wanted to. Thus, most students worry about getting a grade that is good for them but not necessary "A". Consequently, the interpretation of the results from variable an24 is ambiguous. Further, the nonstatistically significant correlation between variables an24 and se13 may indicate ambiguity in the students' assessment of the statement: "I worry that I will not be able to get an "A" in my mathematics course". Therefore, we suggest the removal of variable an24 from EFA, but the observation of students' vague interpretation of this statement contributes to an understanding of students' mathematical self-beliefs. This finding would not have been possible without an in-depth analysis of the variables.

3.2. High Intercorrelations between Variables. The corresponding determinant of the correlation matrix is much less than the critical value of the abovementioned heuristics (0.00001), which indicates severe problems with multicollinearity. Therefore, the correlation matrix was scanned for very high intercorrelations. Several problematic variables were identified, especially, an24, se21, an22, an17, an2, an8, and an26.

Table 3 shows the correlation matrix with only the correlations higher than 0.60. Several pairs of variables were identified having very high intercorrelation coefficients, e.g., an26 and an2 ($r = 0.85$), an22 and se21 ($r = -0.82$), an22 and an17 ($r = 0.79$), and an2 and an8 ($r = 0.75$), respectively. However, to make adequate decisions about the removal of variables required an in-depth analysis.

Variable an26 correlates highly with both an2 and an8, and therefore, the straightforward procedure should be to remove all three variables, except for one. But, since all these variables are associated with a specific dimension of mathematics anxiety: test anxiety, the removal of two of them might result in a too scarce description of mathematics anxiety. Concerning variables an2 and an8, the students seem to have interpreted the variables an2 and an8 in the same way, as asking for the same thing, i.e., an indication of redundancy variables. When students are tense in the preparation of mathematics tests, it is plausible that the students are also worried about taking the mathematics test. Besides, the high intercorrelation between an2 and an26 shows that if a student is nervous or tense to prepare for tests, he/she is probably also worried about taking the test. These noted issues in our data indicate a problem with

TABLE 3: The correlation matrix showing only correlations higher than 0.60.

	an2	se4	an5	an6	se7	an8	se9	se10	se12	se13	an14	an15	se16	an17	an18	se19	se20	se21	an22	se23	an25	an26	an27	se28	se29
an2						0.75‡																**0.85‡**		−0.70‡	
se4						−0.63‡	0.64‡			0.66‡			0.72‡	0.62‡		0.72‡	0.60‡	0.70‡	−0.63‡						−0.62‡
an5						0.61‡												−0.60‡				0.69‡			
an6						0.69‡		−0.68‡					−0.64‡	0.63‡	0.66‡	−0.62‡		−0.77‡	0.66‡		0.73‡				
se7									0.70‡						−0.70‡										
an8							−0.67‡				0.60‡		−0.68‡	0.67‡	0.61‡	−0.80‡	−0.63‡	−0.74‡	0.72‡			0.78‡		−0.62‡	−0.68‡
se9													0.64‡		−0.61‡	0.67‡	0.62‡	0.68‡	−0.63‡					0.75‡	0.60‡
se10									0.63‡								0.66‡								0.61‡
se12														−0.67‡											
se13														−0.62‡	−0.61‡	0.72‡		0.62‡	−0.66‡						
an14																							0.70‡		−0.62‡
an15															0.63‡										
se16																		0.70‡							
an17															0.68‡	−0.70‡	−0.68‡	−0.72‡	−0.68‡	0.61‡					
an18																	−0.65‡	−0.67‡	0.79‡						
se19																	0.63‡	0.76‡	0.71‡			−0.67‡		0.61‡	0.61‡
se20																		0.65‡	−0.68‡					0.64‡	0.69‡
se21																			**−0.82‡**	0.60‡	−0.76‡			0.61‡	
an22																					0.63‡				
se23																									
an25																									
an26																									−0.63‡
an27																									
se28																									0.62‡
se29																									

Note. $N = 79$. Correlations over 0.80 appear in bold. Significance levels ‡$p < 0.001$, †$p < 0.01$, and *$p < 0.05$ (two-tailed).

multicollinearity. Although we are risking a too scarce description of test anxiety, we suggest the removal of two of the three variables (an2, an8, and an26), i.e., the variables an2 and an8.

Our data show that students who are tense in the preparation of a mathematics test are more likely to be worried about taking the mathematics test, which shows an association between the perceived mathematics anxiety during the preparation for a mathematics test and the actual taking of the mathematics test. Besides, this might give some clues for future studies about the sources of students' test anxiety. Consequently, removing very high correlating variables without a profound analysis would result in a loss of information about student's test anxiety.

Further, concerning variable an22, more than half of the students responded that they sometimes or more often worry that they will not be able to understand mathematics. In light of this, an interpretation of the collinearity between variables an22 and an17 ($r = 0.79$) might be that if a student worries about being able to understand mathematics, it means that the student is probably also worried about whether he or she knows enough mathematics to succeed with future courses. This observation might indicate that variables an22 and an17 are redundant variables. Moreover, students could report being less worried about doing well in future courses if they are studying their last mandatory course in mathematics since they have decided not to attend them. This interpretation can be an indication of the low validity of variable an17. Therefore, together with the fact that an17 correlates highly with an22, we suggest the removal of variable an17 from EFA.

A similar validity problem, as mentioned above, concerns variable se21: "I feel that I will be able to do well in future mathematics courses," which also has a high intercorrelation coefficient with variable se19 ($r = 0.76$). Less than half of the students reported that they sometimes or more often feel like they can do well in future mathematical courses. Besides, variables se21 and an22 have a high intercorrelation coefficient ($r = -0.82$), which also supports the interpretation of the variables an17 and an22 as redundancy variables. Hence, we suggest that also variable se21 should be removed from EFA.

However, even if the correlation is high, we sometimes conclude that no variables should be removed. Let us consider, for instance, se28 and se9 ($r = 0.75$). Additionally, the correlation between se9 and se28 is positive. Now, most of the students reported that they never or seldom believe they could think like a mathematician, and they never or seldom consider themselves a person who is good at mathematics. These two items represent different dimensions of an underlying construct because a person who believes in him/herself as good at mathematics does not automatically mean that the same person also thinks she/he can think like a mathematician. The variables se9 and se28 are all related to general mathematics self-efficacy and measure different dimensions of an underlying construct. Although the intercorrelation coefficient is high, we conclude that the variables should not be removed from EFA. This high correlation rather indicates that students believe
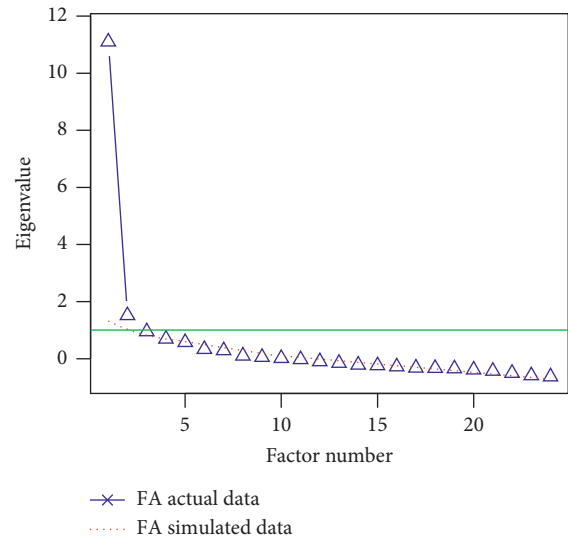

FIGURE 1: Scree plot of eigenvalues for actual and simulated data.

TABLE 4: Parallel analysis results (eigenvalues for actual and simulated data).

Eigenvalue #	Actual	Simulated
1	11.096	1.409
2	1.520	1.003
3	0.950	0.846
4	0.695	0.714
5	0.575	0.595
6	0.333	0.502

that a person who is good at mathematics is a person who also can think like a mathematician. Consequently, students possibly have a preconception about what it takes to be good at mathematics, and an indication of students' static beliefs about intelligence [34], and could thus be a symptom of students' helplessness towards getting a high grade in the mathematics course.

3.3. The Final Factor Model.
For the reader's interest, we also report shortly which factor model our EFA resulted in, and how this model was found, given the data collected for the pilot study.

First, the identified problematic variables an24, an17, se21, an2, and an8 were removed from EFA. After that, the Kaiser–Meyer–Olkin (KMO) measure verified the sampling adequacy for factor analysis KMO = 0.86, which is "great" according to the literature. All the KMO values for the items were over 0.74, which is well above the acceptable minimum of 0.50 [35, 36]. Further, Bartlett's test of sphericity, $\chi^2(276) = 1176.904$, $p < 0.001$, showed that correlations were large enough for EFA. For this study, parallel analysis and scree plot were used for factor retention.

The scree plot (see Figure 1) shows some ambiguity in the interpretation of how many factors to retain, however, based on the parallel analysis (see Table 4); four factors were retained. Since at the fourth simulation, the eigenvalue calculated from the actual data is below the eigenvalue

TABLE 5: The pattern matrix and the summary of the final factor solution.

| Variable | Oblimin rotated factor loadings | | | | h2 | u2 |
	Math self-efficacy	Assignment	In-class anxiety	Test anxiety		
se1	−0.14	0.28	**−0.69**	0.08	0.54	0.46
se4	**0.64**	0.24	−0.16	0.07	0.70	0.30
an5	0.12	−0.33	0.24	**0.63**	0.70	0.30
an6	−0.23	−0.42	0.20	0.20	0.60	0.40
se7	0.01	**0.77**	−0.05	0.02	0.62	0.38
se9	**0.60**	0.12	−0.14	−0.16	0.69	0.31
se10	0.08	**0.59**	−0.12	−0.22	0.63	0.37
se12	0.11	**0.83**	−0.04	0.14	0.76	0.24
se13	**0.82**	−0.05	0.02	0.07	0.57	0.43
an14	−0.19	0.07	**0.77**	0.03	0.72	0.28
an15	−0.20	−0.30	−0.29	**0.55**	0.57	0.43
se16	**0.67**	0.26	−0.01	0.11	0.64	0.36
an18	−0.18	**−0.64**	−0.13	0.28	0.72	0.28
se19	**0.58**	0.06	−0.13	−0.29	0.72	0.28
se20	0.25	0.37	−0.20	−0.28	0.65	0.35
an22	−0.42	−0.39	−0.02	0.25	0.71	0.29
se23	0.53	0.29	−0.13	0.24	0.49	0.51
an25	−0.20	−0.27	0.20	0.30	0.50	0.50
an26	−0.24	0.20	0.26	**0.71**	0.86	0.14
an27	−0.01	0.07	**0.81**	0.08	0.68	0.32
se28	**0.74**	−0.11	0.01	−0.24	0.68	0.32
se29	0.23	0.17	−0.37	−0.31	0.64	0.36
SS loadings	4.82	4.10	2.80	2.69		
% of variance	21.89	18.62	12.72	12.25		
% of explained	33.43	28.44	19.43	18.70		
Cronbach's α	0.89	0.84	0.77	0.74		
Math self-efficacy	1.00	0.54	−0.37	−0.48		
Assignment	0.54	1.00	−0.28	−0.26		
In-class anxiety	−0.37	−0.28	1.00	0.34		
Test anxiety	−0.48	−0.26	0.34	1.00		

Note: factor loadings over 0.55 appear in bold. h2 = communalities; u2 = uniqueness.

TABLE 6: The structure matrix.

| Variable	Math self-efficacy	Assignment	In-class anxiety	Test anxiety
se1	0.23	0.37	−0.69	−0.16
se4	0.79	0.61	−0.44	−0.36
an5	−0.45	−0.50	0.50	0.74
an6	−0.62	−0.65	0.47	0.49
se7	0.43	0.79	−0.26	−0.20
se9	0.80	0.53	−0.45	−0.53
se10	0.55	0.73	−0.39	−0.46
se12	0.51	0.86	−0.26	−0.15
se13	0.75	0.37	−0.25	−0.31
an14	−0.45	−0.25	0.83	0.37
an15	−0.52	−0.47	0.06	0.62
se16	0.76	0.60	−0.29	−0.29
an18	−0.61	−0.77	0.21	0.49
se19	0.79	0.48	−0.46	−0.62
se20	0.66	0.63	−0.49	−0.56
an22	−0.74	−0.68	0.33	0.55
se23	0.61	0.55	−0.32	−0.13
an25	−0.56	−0.51	0.45	0.53
an26	−0.57	−0.19	0.54	0.86
an27	−0.31	−0.18	0.82	0.34
se28	0.79	0.35	−0.31	−0.56
se29	0.61	0.48	−0.62	−0.60
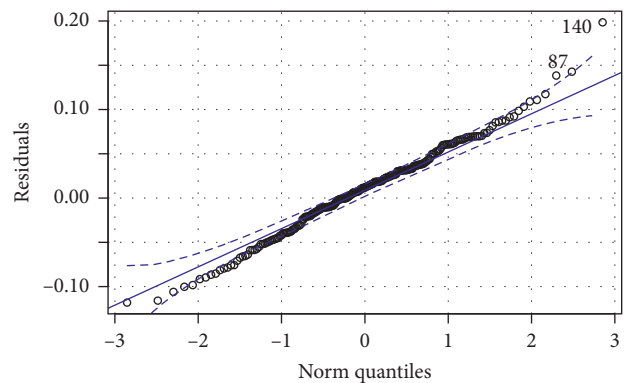
FIGURE 2: Q-Q plot of factor residuals with 95% confidence levels.

calculated from the simulated data. Further, since these factors are very likely to be correlated because of the strong relationship between mathematics self-efficacy and anxiety towards mathematics, the oblique rotation was used.

Furthermore, during the factor analysis, two variables were excluded: an11 and an3 because of low communality (0.40) (c.f. [8]). The significant loading cutoff was 0.55 (cf. [4, 10]). The final factor model presented in Table 5 was produced using the minimum residuals method with

oblique rotation (oblimin). The corresponding structure matrix is presented in Table 6. The retained four factors in the final factor model could explain 65% of the variance (see Table 5).

A measure of factor model fit is the sum of the squared residuals divided by the sum of the correlations, which for this factor model is 0.99 (values over 0.95 indicate a good fit [4]). Besides, fewer than 50% of the factor residuals have absolute values greater than 0.05, and the factor residuals are approximately normally distributed (see Figure 2). All these measures indicate a good factor model fit [4].

The items that load on the same factor suggest that factor 1 represents general mathematics self-efficacy, factor 2 abilities and anxieties to complete assignments, factor 3 anxiety towards mathematics in classroom settings, and factor 4 anxiety towards evaluations, which we label test anxiety. The first two subscales had high reliabilities with $\alpha = 0.89$ and $\alpha = 0.84$, and in-class anxiety and test anxiety subscales had $\alpha = 0.77$ and $\alpha = 0.74$.

4. Conclusions

The qualitative criteria that we used above for removing items from EFA were based upon the indications of redundancy or low validity. Redundancy concerns variables that are interpreted in the same way by the respondents, and a variable with low validity does not seem to measure what it was intended to measure. For example, variable an24, "I worry that I will not be able to get an "A" in my mathematics course," was removed from EFA because students might assess the item differently depending on which grade they have set as a goal to themselves. Hence, we concluded that the item has low validity. Further, variable an24 has lots of correlation coefficients below 0.30. Therefore, if we instead had decided to keep variable an24, it probably would have trouble loading into the latent variables.

Our findings show that significant information can be reliably measured with a lesser number of items. For example, although two problematic variables related to test anxiety were removed from EFA, test anxiety appeared as a meaningful factor with some unexpected but interpretable loadings in the final factor model. The item "I get nervous when taking a mathematics test" has the strongest loading in the test anxiety factor, but the item "I worry that I will not be able to use mathematics in my future career when needed" also loads on the same factor and has to do with situations where students mathematics knowledge is evaluated. Also, the item "working on mathematics homework is stressful for me" loaded on this factor and this may be interpreted so that students mainly work on mathematics homework when preparing themselves for an exam.

Although in this study we have presented an adequate EFA based on a small sample size, we acknowledge that large sample sizes most often contribute to better generalizability [8]. Concerning the present data, a more relevant question is what makes a sample size adequate for analysis. The answer seems to depend on how many variables represent each factor and the strength of variable loadings and communalities. By the literature we have referred to, data must be strong, i.e., every factor has to be related to, at least, four items with high loadings and high communalities and there should not be any cross-loadings [8, 22]. All extracted factors in the final factor model above have items with high loadings. However, only the first two factors have at least four items with high loadings and communalities. Although there are some nonsignificant "cross-loadings" in the factor model, data can be considered strong in light of typical acceptable factor models.

Multicollinearity is common in data that have been collected for studies in social sciences. In most cases, it is not taken as a big issue. Researchers simply have to cope with some degree of multicollinearity. One way of decreasing the degree of multicollinearity is to increase sample size, which in most cases decreases the standard errors. Still, to increase the validity of discovered latent variables, it is important to assess the severity of the multicollinearity and making a profound analysis of problematic variables. Regardless of sample size, extreme multicollinearity is problematic and could in the worst cases result in an ambiguous factor model (c.f. [6].)

A thorough quantitative study often includes the usage of EFA or another method suitable for constructing sum variables that represent the latent variables. However, before conducting an EFA, a pretreatment of data is required to make data more adequate for EFA. That includes a profound analysis of the correlations between variables and possible removal of problematic variables. This procedure can be a complex process and, in many cases, requires decisions that have to be taken on subjective criteria, for example: What is the adequate sample size for EFA? What quantitative criteria are suitable for identifying variables that are the source of extreme multicollinearity? We showed in this study that there are no general objective answers to these questions, and we emphasize the importance of analyzing data qualitatively when we make this type of subjective decision.

Data Availability

The anonymised version of data used in this study is available upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] L. Bonne and M. Johnston, "Students' beliefs about themselves as mathematics learners," *Thinking Skills and Creativity*, vol. 20, pp. 17–28, 2016.

[2] E. M. Skaalvik and S. Skaalvik, "Self-concept and self-efficacy: a test of the internal/external frame of reference model and predictions of subsequent motivation and achievement," *Psychological Reports*, vol. 95, no. 3, pp. 1187–1202, 2004.

[3] T. Tossavainen, R. J. Rensaa, and M. Johansson, "Swedish first-year engineering students' views of mathematics, self-efficacy and motivation and their effect on task performance,"

International Journal of Mathematical Education in Science and Technology, p. 1, 2019.

[4] A. Field, J. Miles, and Z. Field, Discovering Statistics Using R, SAGE Publications, Oaks, CF, USA, 2013.

[5] L. Cohen, L. Manion, and K. Morrison, Research Methods in Education, Routledge, Milton Park, AB, USA, 8 edition, 2017.

[6] H. W. Marsh, M. Dowson, J. Pietsch, and R. Walker, "Why multicollinearity matters: a reexamination of relations between self-efficacy, self-concept, and achievement," Journal of Educational Psychology, vol. 96, no. 3, pp. 518–522, 2004.

[7] A. J. Martin, H. W. Marsh, and R. L. Debus, "Self-handicapping and defensive pessimism: a model of self-protection from a longitudinal perspective," Contemporary Educational Psychology, vol. 28, no. 1, pp. 1–36, 2003.

[8] A. B. Costello and J. W. Osborne, "Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis," Practical Assessment, Research and Evaluation, vol. 10, no. 7, 2005.

[9] P. Samuels, "Advice on reliability analysis with small samples," Techical Report, Birmingham City University, London, UK, 2015.

[10] A. G. Yong and S. Pearce, "A beginner's guide to factor analysis: focusing on exploratory factor analysis," Tutorials in Quantitative Methods for Psychology, vol. 9, no. 2, pp. 79–94, 2013.

[11] B. G. Tabachnick and L. S. Fidell, Using Multivariate Statistics: Pearson New International Edition, Pearson Education Limited, London, UK, 2013.

[12] R. C. Rockwell, "Assessment of multicollinearity," Sociological Methods & Research, vol. 3, no. 3, pp. 308–320, 1975.

[13] P. Y. Chen, M. Smithson, and P. M. Popovich, Correlation: Parametric and Nonparametric Measures, SAGE Publications, Oaks, CF, USA, 2002.

[14] J. Choi, M. Peters, and R. O. Mueller, "Correlational analysis of ordinal data: from pearson's r to bayesian polychoric correlation," Asia Pacific Education Review, vol. 11, no. 4, pp. 459–466, 2010.

[15] V. Savalei, "What to do about zero frequency cells when estimating polychoric correlations," Structural Equation Modeling: A Multidisciplinary Journal, vol. 18, no. 2, pp. 253–273, 2011.

[16] M. Rhemtulla, P. É. Brosseau-Liard, and V. Savalei, "When can categorical variables be treated as continuous? a comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions," Psychological Methods, vol. 17, no. 3, pp. 354–373, 2012.

[17] G. Norman, "Likert scales, levels of measurement and the "laws" of statistics," Advances in Health Sciences Education, vol. 15, no. 5, pp. 625–632, 2010.

[18] J. C. F. de Winter, D. Dodou, and P. A. Wieringa, "Exploratory factor Analysis with small sample sizes," Multivariate Behavioral Research, vol. 44, no. 2, pp. 147–181, 2009.

[19] D. J. Mundfrom, D. G. Shaw, and T. L. Ke, "Minimum sample size recommendations for conducting factor analyses," International Journal of Testing, vol. 5, no. 2, pp. 159–168, 2005.

[20] R. C. MacCallum, K. F. Widaman, S. Zhang, and S. Hong, "Sample size in factor analysis," Psychological Methods, vol. 4, no. 1, pp. 84–99, 1999.

[21] R. C. MacCallum, K. F. Widaman, K. J. Preacher, and S. Hong, "Sample size in factor analysis: the role of model error," Multivariate Behavioral Research, vol. 36, no. 4, pp. 611–637, 2001.

[22] B. Thompson, Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications, American Psychological Association, Washington, DC, USA, 2004.

[23] A. Bandura, "Self-efficacy mechanism in human agency," American Psychologist, vol. 37, no. 2, pp. 122–147, 1982.

[24] M. Bong and E. M. Skaalvik, "Academic self-concept and self-efficacy: how different are they really?" Educational Psychology Review, vol. 15, no. 1, pp. 1–40, 2003.

[25] B. J. Zimmerman, "Self-efficacy: an essential motive to learn," Contemporary Educational Psychology, vol. 25, no. 1, pp. 82–91, 2000.

[26] D. K. May, Mathematics self-efficacy and anxiety questionnaire, Ph.D. thesis, University of Georgia, Athens, GA, USA, 2009.

[27] E. D. Peña, "Lost in translation: methodological considerations in cross-cultural research," Child Development, vol. 78, no. 4, pp. 1255–1264, 2007.

[28] Y. Rosseel, "lavaan: an R package for structural equation modeling," Journal of Statistical Software, vol. 48, no. 2, pp. 1–36, 2012.

[29] R Core Team, R: A Language and Environment for Statistical ComputingRC Team, Vienna, Austria, 2020, http://www.R-project.org/.

[30] L. T. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives," Structural Equation Modeling: A Multidisciplinary Journal, vol. 6, no. 1, pp. 1–55, 1999.

[31] T. A. Brown, "Methodology in the social sciences," Confirmatory Factor Analysis for Applied Research, Guilford Press, New York, NY, USA, 2 edition, 2015.

[32] S. van Buuren and K. Groothuis-Oudshoorn, "Mice: multivariate imputation by chained equations in R," Journal of Statistical Software, vol. 45, no. 3, pp. 1–67, 2011, https://www.jstatsoft.org/v45/i03/.

[33] W. Revelle, Psych: Procedures for Psychological, Psychometric, and Personality ResearchEvanston, Illinois, IL, USA, 2020, https://CRAN.R-project.org/package=psych.

[34] C. S. Dweck and E. L. Leggett, "A social-cognitive approach to motivation and personality," Psychological Review, vol. 95, no. 2, pp. 256–273, 1988.

[35] H. F. Kaiser, "An index of factorial simplicity," Psychometrika, vol. 39, no. 1, pp. 31–36, 1974.

[36] H. F. Kaiser and J. Rice, "Little jiffy, mark iv," Educational and Psychological Measurement, vol. 34, no. 1, pp. 111–117, 1974.