

Article

Not peer-reviewed version

Validity evidence for using the Situational Motivation Scale to assess pre-clerkship medical student motivation.

Brian Wasicek and [Douglas McHugh](#)*

Posted Date: 23 May 2024

doi: 10.20944/preprints202405.1525.v1

Keywords: Medical students; motivation; situational motivation scale; SIMS validity evidence; undergraduate medical education; pre-clerkship curriculum



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Validity Evidence for Using the Situational Motivation Scale to Assess Pre-Clerkship Medical Student Motivation

Brian Wasicek ^{1,2} and Douglas McHugh ^{1,*}

¹ Frank H. Netter MD School of Medicine, Quinnipiac University, Hamden, CT 06518, USA; bwasicek@uchc.edu

² The University of Connecticut Health Center, Farmington, CT 06030, USA

* Correspondence: douglas.mchugh@quinnipiac.edu

Abstract: Motivation is essential in education, with highly motivated learners engaging more deeply with content and more ably transferring knowledge to new contexts. However, the validity of scales to measure motivation has been underexplored in pre-clerkship medical education. This study evaluates the validity of the Situational Motivation Scale (SIMS) for measuring motivation among pre-clerkship medical students in post-situational and short-term contexts. Using a sample of $n = 156$ students from the Frank H. Netter MD School of Medicine at Quinnipiac University, SIMS was tested for construct validity with attention paid to content, response process, internal structure, relationships to other variables, and consequences of use evidence. Small modifications from present to past tense in English were made for clarity following focus group feedback, and content validity was ensured via expert consultation. SIMS demonstrated strong internal consistency with a satisfactory Cronbach's alpha for all subscales and anticipated patterns of correlations. Factor analysis confirmed appropriate factor loadings, with a stronger model fit for the short-term context, and no observed adverse effects on student engagement. These findings support the robustness of SIMS in capturing intrinsic, extrinsic, and amotivation in pre-clerkship medical students, highlighting its applicability for short-term and situational motivational assessment.

Keywords: medical students; motivation; situational motivation scale; SIMS validity evidence; undergraduate medical education; pre-clerkship curriculum

1. Introduction

Medical school begins the challenging journey of equipping students for the demanding profession of being a physician. A career where the stakes are high – involving people's lives and well-being. As such it requires discipline, perseverance, and motivation. The latter is crucial for medical students because of the rigorous and demanding nature of medical school curricula. Put simply, motivation is the general desire or willingness of someone to do something [1]. It energizes, directs and sustains behavior. Self-Determination Theory (SDT) posits that motivation is the fundamental reason underlying individuals' thoughts, actions, and behaviors. [2–4]. Motivation is of particular concern in education and is something that should be paid attention to from a learning design and curriculum evaluation perspective because when learners are motivated they try harder to understand, they process information at a deeper level, and they are more apt to apply what they have learned to novel situations [5]. In the context of well-being and burnout, students embarking on their medical school journey in the United States initially exhibit lower levels of burnout, diminished rates of depression, and an enhanced quality of life compared to their peers graduating from college into different professions. However, this trend inverts by the second year of medical school. At this point, burnout—characterized by emotional exhaustion, heightened stress, a reduced sense of

personal achievement, and thoughts of suicide—intensifies, reaching its zenith during the residency period [6–8]. The expectations placed upon medical students are continually increasing, which amplifies the imperative to support them in developing self-help strategies [9–12]. Being able to authentically measure and track motivation should assist medical educators and curriculum designers with this.

How can we measure motivation? To answer that question, we have to appreciate that motivation is a construct. Constructs are conceptual ideas or abstractions used to facilitate an understanding of something intangible [13]. All scientific fields are founded upon frameworks consisting of constructs and the interplay between them. The natural sciences employ constructs like gravity, phylogenetic dominance, and tectonic pressure. Similarly, the behavioral sciences utilize constructs including conscientiousness, sportsmanship, and resilience. Motivation, like these, is a real albeit abstract human experience. Researchers often use questionnaires with multiple items to measure behavioral constructs. Item responses are added up to a score and it is hypothesized that this score represents a person's position on the construct. Behavioral constructs differ in the ease with which they can be measured. Some are straightforward enough to be evaluated with just a single question or a few inquiries. In contrast, others are multifaceted and complex, comprising various dimensions unified by a central theme that collectively defines the construct. Such constructs may necessitate an extensive series of questions to be fully operationalized [14].

Motivation is a complex construct that contains four dimensions: intrinsic motivation, identified regulation, external regulation, and amotivation [2]. SDT perceives these motivational dimensions as existing on a continuum, ranging from self-directed (autonomous) to those driven by external forces [2–4] (Figure 1). In the context of SDT, "regulation" refers to the processes that govern why and how a person initiates and maintains a particular behavior over time, especially in the face of challenges and obstacles. *Intrinsically motivated* students learn because they derive pleasure and satisfaction, or get a sense of accomplishment, from learning. They enjoy engaging in the behavior for its own sake (e.g., "I enjoy running, it makes me feel good"). *Identified regulation* happens when an individual acknowledges the value and importance of a behavior, even if it does not provide enjoyment or immediate gratification (e.g., "I run because I know running is good for me"). *External regulation* arises when behavior is motivated by the pursuit of rewards or the avoidance of adverse outcomes, with the learner feeling compelled to act in a certain manner due to external pressures (e.g., "I run because I will be given a race medal"). *Amotivation* represents a total lack of motivation, characterized by a learner's reluctance to engage or exert effort, often accompanied by an absence of purpose or anticipation of reward. This state is akin to learned helplessness, where individuals are beset by a sense of ineffectuality and perceive no opportunity to alter the trajectory of events (e.g., "I've stopped running. I can't see the point anymore – it just tires me out"). These four dimensions can and do vary from person to person over time. Educators aim for medical students to be guided by intrinsic motivation and identified regulation instead of relying on external rewards and punishments or succumbing to a lack of motivation.

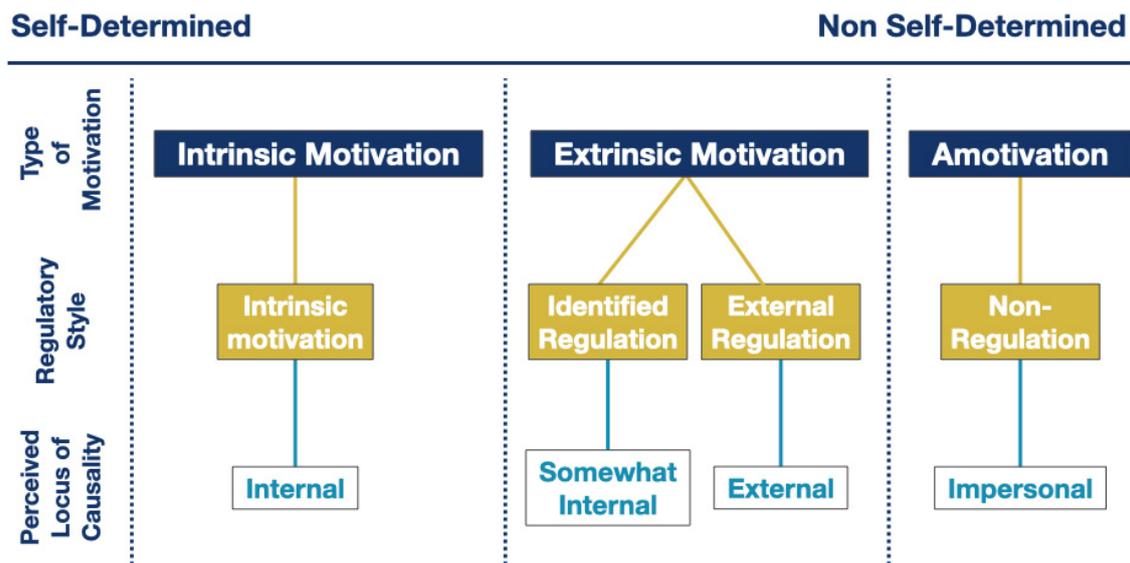


Figure 1. Visual representation of the continuum of motivation from self-determined to non self-determined according to self-determination theory [2].

The Situational Motivation Scale (SIMS) is an instrument informed by self-determination theory (SDT) to measure motivation [15]. It was developed by Frederic Guay and colleagues in 2000 who published a series of studies that support its use within a French-Canadian college context (in Quebec students may enroll in a two-year college program between high school and university) [15], and has since been used by many researchers to study academic motivation including within the biomedical sciences [16]. It consists of 16 questions (Appendix A). Four for each of the four dimensions that make up the construct of motivation. It was designed to be used situationally; to capture a measure of student motivation in situ at one point in time. In the United States and Canada, medical school operates as a graduate-level program. Pre-clerkship curricula often blend classroom instruction in integrated basic sciences with early practical exposure to real patients and the practice of medicine. Therefore, pre-clerkship medical students have to sustain a desire or willingness (i.e., be motivated) to participate authentically during learning events as well as to study further and pursue mastery after learning events conclude [17]. These two sets of circumstances represent two different but related contexts; the former 'situational' and the latter 'short-term'. However, interrupting educational events to have learners complete in situ the 16-items that make up SIMS would be disruptive and impractical for educators or curriculum designers to sustain on a routine basis; Guay et al. interrupted the French-Canadian students while they were studying alone in a college library. A 'post-situational' assessment (i.e., proximal to the conclusion of a learning event) is more realistic for medical students. We define the *post-situational* context to mean within an hour of a learning event ending and the *short-term* context to mean later that day or within a week of a learning event ending. Since measuring motivation in these contexts represents a novel use of the SIMS instrument, fresh validity evidence to support it must be obtained. Messick (1989) defined validity as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores and other modes of assessment." [18]. In other words, validity refers to data presented to support or refute the interpretation assigned to SIMS scores [19]. Without evidence of validity, the SIMS scores obtained from medical student learners have little or no intrinsic meaning.

Thus, the objective of this study was to answer the question of whether SIMS is supported by validity evidence to measure the true state of affairs with regard to learner motivation in medical student pre-clerkship post-situational and short-term contexts.

2. Materials and Methods

2.1. Conceptual Approach to Construct Validity Evidence

Current perspectives advocate that all forms of validity be unified within a singular concept of 'construct validity', which refers to the extent to which a score accurately reflects the intended underlying construct [20,21]. Messick recognized five pillars of evidence that underpin construct validity: content, response processes, internal structure, relationships to other variables, and the implications of assessment outcomes (i.e., consequences) [18] (Table 1). It is important to note that, "these are not different types of validity but rather they are categories of evidence that can be collected to support the construct validity of inferences made from instrument scores" [20]. We followed this approach as a scheme to collect and evaluate psychometric data pertaining to SIMS in post-situational and short-term medical student contexts.

Table 1. Five sources of construct validity evidence and their definitions [19].

Validity Evidence Sources	Validity Evidence Definitions
Content	Content validity focuses on the construct's dimensions and how well they are represented.
Response Process	Response process validity is the extent to which the actions and thought processes of respondents demonstrate that they understand the instrument questions the same way as the researchers do.
Internal Structure	Internal structure evidence asks whether a group of questions designed to assess the same dimension exhibit higher correlations with one another compared to groups of questions aimed at measuring different dimensions.
Relationship to Other Variables	Relationship to other variables evidence asks if external correlations (i.e., correlations to entirely different instruments) are confirmatory or counter-confirmatory.
Consequences	Consequences evidence asks what is the impact on respondents of being appraised by the instrument?

Content validity focuses on the construct's dimensions and how well they are represented [19]. *Response process* validity is the extent to which the actions and thought processes of respondents demonstrate that they understand the instrument questions, and how to go about answering them, the same way as the researchers do [19]. It seeks to identify and eliminate sources of error related to the instrument and how it is administered. Evidence of *internal structure* seeks to determine whether a group of questions designed to assess the same dimension exhibit higher correlations with one another compared to groups of questions aimed at measuring different dimensions [19]. In other words, do the 4 questions that are trying to measure intrinsic motivation 'hang together' better with each other than they do with the other 3 question sets that aren't trying to measure intrinsic motivation? *Relationship to other variables* evidence asks if external correlations (i.e., correlations to entirely different instruments) are confirmatory or counter-confirmatory [19]? And *consequences* evidence asks what is the impact on respondents of being appraised by the instrument [19]? Could someone who agrees to answer the 16 questions be harmed or substantially benefit in anyway? It seeks to identify whether the appeal of any benefits or the fear of any harms could introduce sources of bias into the data collected.

2.2. Study Design and Sample

This single institution validation study invited n = 186 first- and second-year medical students from the Frank H. Netter MD School of Medicine at Quinnipiac University to participate. Participation in the study was voluntary and de-identified.

2.3. Procedures

All participants were asked to what extent they agreed on a 7-point scale (1 = not at all; 7 = completely agree) with the items on the following four survey instruments: Situational Motivation Scale (SIMS), Academic Motivation Scale (AMS), Multidimensional Crowdsworker Motivation Scale (MCMS), and Maslach Burnout Inventory Student Survey Exhaustion component (MBI-SSEE) [15,22–24]. The *post-situation* and *short-term* versions of these surveys were administered on two separate occasions to all 156 study participants (Appendixes B and C). The 83 first-year medical students completed the four surveys with regard to a two-hour small group learning event (each small group comprised 8 students + 1 faculty facilitator) on the central nervous system's sensory systems (i.e., somatosensory, visual, auditory, olfactory, and gustatory) from a post-situational and short-term perspective; whereas, the 73 second-year medical students completed them with regard to a two-hour small group learning event (each small group comprised 8 students + 1 faculty facilitator) on pathology of kidney disorders from a post-situational and short-term perspective. All versions of the SIMS, MCMS, AMS, and MBI-SSEE were administered as paper forms. Survey packets were marked with anonymous numerical identification. Participants received verbal instructions before the survey administration commenced. Data were entered independently by authors B.W. and D.M. then compared to ensure data entry was accurate.

2.4. Scale Adaptation

To adapt SIMS to both post-situational and short-term contexts, minor grammatical changes were made to both the initial prompt and to the individual items (Table 2). This was done to convert the language from present to past tense English. For example, the prompt, "Why are you engaged in this activity?" became "Why were you engaged in this activity?"; and item 1 "Because I think that the activity is interesting" became "Because I thought that the activity was interesting."

Table 2. The situational motivation scale (SIMS)'s 16 items, which were adapted from present to past tense [15].

Item	Present Tense SIMS Version	Past Tense SIMS Version
1	Because I think that the activity is interesting	Because I thought that the activity was interesting
2	Because I am doing it for my own good	Because I was doing it for my own good
3	Because I am supposed to do it	Because I was supposed to do it
4	There may be good reasons to do this activity, but personally I don't see any	There may have been good reasons to do this activity, but personally I don't see any
5	Because I think that this activity is pleasant	Because I thought that this activity was pleasant
6	Because I think that this activity is good for me	Because I thought that this activity was good for me
7	Because it is something that I have to do	Because it was something that I had to do
8	I do this activity but I am not sure if it is worth it	I did this activity but I am not sure if it was worth it
9	Because this activity is fun	Because this activity was fun
10	By personal decision	By personal decision
11	Because I don't have any choice	Because I didn't have any choice
12	I don't know; I don't see what this activity brings me	I don't know; I don't see what this activity brought me
13	Because I feel good when doing this activity	Because I felt good when doing this activity
14	Because I believe that this activity is important for me	Because I believe that this activity was important for me
15	Because I feel that I have to do it	Because I feel that I had to do it
16	I do this activity, but I am not sure it is a good thing to pursue it	I did this activity, but I am not sure it was a good thing to pursue it

2.5. Construct Validity – Content Analysis

Four subject matter experts experienced with psychometric data analysis and familiar with self-determination theory (1 PhD statistician and professor of medical sciences, 1 PhD educator and professor of learning sciences, 1 PhD educator and associate dean of medical education, and 1 EdD educator and assistant professor of education) reviewed Guay et al. (2000)'s original content validity evidence and our grammatical edits to determine if they represented substantial content changes [15].

2.6. Construct Validity – Response Process Analysis

Six focus groups were held, with $n = 7$ medical students per session, to review what they understood the 16 SIMS questions to mean and to evaluate the extent to which our proposed survey administration process could be confusing to respondents. This was done in order to maximize response integrity and avoid data collected being skewed or in error because questions were misunderstood or instructions on how to fill out the SIMS survey were unclear or misleading. Standardized survey administration and data collection procedures were finalized at the conclusion of this aspect of the study.

2.7. Construct Validity – Internal Structure Analysis

Internal structure of the SIMS instrument for both post-situational and short-term contexts was assessed via four statistical analyses. The data were analyzed using IBM SPSS Statistics, Version 27 and Mplus v8.2 [25].

Initially, to estimate reliability, Cronbach's alpha was computed for each of the four dimensions of the motivation construct. Cronbach's alpha assesses internal consistency by indicating the extent to which a set of items is interrelated. A 'good' coefficient of Cronbach's alpha is generally one that is substantially high, implying that the items consistently measure the same underlying concept or construct. Methodologists frequently suggest a threshold of 0.7 or above as an acceptable value [26–29].

Secondly, we assessed internal consistency through the calculation of Pearson's r , analyzing the patterns of correlations. Given that the four dimensions of motivation are situated sequentially along a continuum, the interrelationships among the four item sets of the SIMS should display a simplex pattern. This pattern dictates that items closer on the motivational continuum will exhibit stronger positive correlations than those further apart.

Thirdly, to affirm the configurational structure of the 16-item SIMS scale, confirmatory factor analysis (CFA) was conducted utilizing Mplus software. Factor is an alternative term for dimension. Following the statistical analyses reported by Guay (2000) [15], each subset of four questions was loaded onto each of the four factors and then Geomin rotation allowed the factors to correlate [30]. Model goodness-of-fit was assessed by The adequacy of the model's fit was evaluated using several indices: the Bayesian Information Criterion (BIC), the ratio of the chi-square to degrees of freedom (χ^2/df), the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR) indices [31]. A χ^2/df value of < 2.0 [32], a CFI value > 0.95 [33]; a RMSEA value < 0.06 [33]; and a SRMR value of ≤ 0.08 [33] were deemed to be indicative of an excellent model fit. The BIC is a criterion for model selection among a set of models [32].

Fourthly, the Wald test assessed the significance of parameter estimates and modification index (MI) values assessed how a model could be improved by allowing additional paths to be estimated. Modification indices help us answer 'what if?' questions about how to improve construct models. Values bigger than 3.84 ($p < 0.05$) or 10.83 ($p < 0.01$) indicate how the model could be improved [34].

2.8. Construct Validity – Relationship to Other Variables Analysis

Relationship to other variables was assessed by calculating the Pearson's r for each combination of factors within the SIMS and the other three instruments (AMS, MCMS, MBI-SSEE). The patterns of correlational relationships between the theoretical constructs measured in each instrument were

then examined for predicted simplex-like patterns. Data analysis was performed using IBM SPSS Statistics, Version 27.

2.9. Construct Validity – Relationship to Other Variables Analysis

The subject matter experts and focus group participants were asked if they considered a respondent completing or declining to complete the SIMS instrument to have any negative repercussions or to possibly harm learners.

2.10. Ethical Approval

The study received approval from the Quinnipiac University Institutional Review Board (protocol # 06617). Informed consent was acquired from all participants in the study.

3. Results

3.1. Participants

Of the 186 medical students invited, 156 agreed (84% response rate): 83 first-year students (57% female, 43% male, mean age = 25) and 73 second-year students (49% female, 51% male, mean age = 26). The data from five participants were excluded from the analysis because their surveys were either incomplete, improperly completed (e.g., circling multiple response numbers for a single item), or displayed straightlining (e.g., consistently circling the same response number for all items across all instruments).

3.2. Construct Validity – Content Evidence

The four subject matter experts concurred that the subtle linguistic modifications from present to past tense did not negatively impact the substance of the SIMS instrument. They affirmed that the items remained appropriately aligned with the conceptual definitions of intrinsic motivation, identified regulation, external regulation, and amotivation.

3.3. Construct Validity – Response Process Evidence

During the focus groups, survey wording and instructions were examined for potential confusion or misunderstandings. Focus group participants were unanimously comfortable with the format of the proposed procedures for survey administration and data collection. They agreed that survey volume was not onerous and data collection was not prone to systematic errors. It was confirmed they understood the 16 SIMS questions to mean the same thing as the study authors B.W. and D.M.

3.4. Construct Validity – Internal Structure Evidence

The internal structure of the SIMS was assessed by examining simplex patterns of internal Pearson's r correlations and reliability as reflected by Cronbach's alpha. SIMS Cronbach's alpha values were 0.98 (IM), 0.98 (IR), 0.95 (ER), and 0.93 (AM) for the post-situational context; and 0.98 (IM), 0.82 (IR), 0.99 (ER), and 0.91 (AM) for the short-term context (Table 3). The internal consistency of SIMS was acceptable given that the lowest Cronbach's alpha value was 0.82, which exceeds the ≥ 0.7 minimum criterion [26–29]. These measures of reliability indicated that questions purporting to measure the same dimension of the motivation construct scored similarly to one another.

Table 3. SIMS Cronbach's alpha values for the four dimensions of the motivation construct calculated from *post-situational* and *short-term* context data.

SIMS Context	Intrinsic Motivation (IM) Qs 1, 5, 9, 13	Identified Regulation (IR) Qs 2, 6, 10, 14	External Regulation (ER) Qs 3, 7, 11, 15	Amotivation (AM) Qs 4, 8, 12, 16
Post-situational Cronbach's alpha	0.98	0.98	0.95	0.93
Short-term Cronbach's alpha	0.98	0.82	0.99	0.91

The internal Pearson's r correlations of SIMS followed an ordered simplex patterns in both short-term and post-situational contexts. Theoretically proximal dimensions (e.g., IM to IR) showed positive correlations, while distal dimensions (e.g., IM to AM) showed negative correlations (Tables 4 and 5).

Table 4. SIMS internal Pearson's r correlation values calculated from the *post-situational* context data form an ordered simplex pattern between the four dimensions of the motivation construct.

SIMS Post-situational Context				
	Intrinsic Motivation (IM)	Identified Regulation (IR)	External Regulation (ER)	Amotivation (AM)
Intrinsic Motivation (IM)	1.0	0.63	-0.25	-0.3
Identified Regulation (IR)	0.63	1.0	-0.16	-0.43
External Regulation (ER)	-0.25	-0.16	1.0	0.3
Amotivation (AM)	-0.3	-0.43	0.3	1.0

Table 5. SIMS internal Pearson's r correlation values calculated from the *short-term* context data form an ordered simplex pattern between the four dimensions of the motivation construct.

SIMS Short-term Context				
	Intrinsic Motivation (IM)	Identified Regulation (IR)	External Regulation (ER)	Amotivation (AM)
Intrinsic Motivation (IM)	1.0	0.27	-0.19	-0.23
Identified Regulation (IR)	0.27	1.0	0.18	-0.29
External Regulation (ER)	-0.19	0.18	1.0	0.16
Amotivation (AM)	-0.23	-0.29	0.16	1.0

Concerning the confirmatory factor analysis (CFA) models for the post-situational and short-term context data, all items demonstrated significant loadings on their intended factors in both models (all p -values < 0.01) (Appendix D). Additionally, the strength of these loadings was consistent across the post-situational and short-term models (Tables 6 and 7).

Table 6. SIMS standardized factor loadings from CFA models of the *post-situational* context data. Note: * $p < 0.01$.

SIMS Post-situational Context				
Item	Intrinsic Motivation (IM)	Identified Regulation (IR)	External Regulation (ER)	Amotivation (AM)
1	0.76*			
5	0.87*			
9	0.85*			
13	0.77*			
2		0.83*		
6		0.85*		
10		0.35*		
14		0.77*		
3			0.76*	
7			0.84*	
11			0.68*	
15			0.85*	
4				0.83*
8				0.78*
12				0.86*
16				0.66*
Factor Correlation		IM	IR	ER
		IR	0.69*	
		ER	-0.25*	-0.07
		AM	-0.30*	-0.48*
				0.28*

Table 7. SIMS standardized factor loadings from CFA models of the *short-term* context data. Note: † = $p < 0.05$; * = $p < 0.01$.

SIMS Short-term Context				
Item	Intrinsic Motivation (IM)	Identified Regulation (IR)	External Regulation (ER)	Amotivation (AM)
1	0.56*			
5	0.83*			
9	0.98*			
13	0.68*			
2		0.60*		

6		0.79*		
10		0.47*		
14		0.79*		
3			0.71*	
7			0.90*	
11			0.50*	
15			0.77*	
4				0.77*
8				0.75*
12				0.75*
16				0.81*
Factor Correlation		IM	IR	ER
		IR	0.25*	
		ER	-0.18†	0.44*
		AM	-0.12	-0.36*

Correlations between factors were mostly in the low to moderate range, supporting the discriminant validity of the factors (Tables 6 and 7). However, the magnitude of the correlations between factors appears to differ between the post-situational and short-term contexts. The post-situational context model showed a strong correlation between IM and IR ($r = 0.69$, $p < 0.01$), while the correlation was smaller for the short-term context model ($r = 0.25$, $p < 0.01$). This is not surprising. A medical student engaging in self-directed study in the short-term (e.g., one week after a learning event) might reasonably be expected to exemplify the overlap of intrinsic motivation—driven by genuine interest and satisfaction in learning medicine—and identified regulation, as they recognize and value the importance of this knowledge for their future role as a competent and effective physician. The AM factor significantly correlated with IM ($r = -0.30$, $p < 0.01$) and ER ($r = 0.28$, $p < 0.01$) in the post-situational context model, but not in the short-term context model (IM: $r = -0.12$, $p = 0.16$; ER: $r = -0.12$, $p = 0.21$). Further, there was a moderately strong positive correlation between ER and IR in the short-term context model ($r = 0.44$, $p < 0.01$), but a small negative correlation in the post-situational context model ($r = -0.07$, $p = 0.47$).

The model goodness-of-fit indices (Tables 8 and 9) indicated the short-term model achieved good fit (BIC = 8025.6, $\chi^2/df = 1.49$, CFI = 0.954, RMSEA = 0.056, SRMR = 0.082), thus supporting the configural structure of the SIMS scale for that context. Whereas the goodness-of-fit indices for the post-situational model suggested room for enhancement in its configurational structure (BIC = 8235.9, $\chi^2/df = 2.5$, CFI = 0.894, RMSEA = 0.099, SRMR = 0.079). In keeping with this there were two MI suggestions to improve model fit for the post-situational context. The first MI (24.7) indicated that having SIMS Q1 (an IM question) reassigned to the IR factor would improve fit. The second MI (10.5) indicated that Q2 (an IR question) might cross-load onto the IM factor. This helps explain the large correlation found between IM and IR in the CFA post-situational context model. The lower BIC value for the short-term model relative to the post-situational model also indicates that SIMS has a better fit for short-term contexts (Tables 8 and 9).

Table 8. SIMS goodness-of-fit indices from CFA models of the *post-situational* context data.

SIMS Post-situational Context		
Model Fit Indices	BIC	8235.9
	χ^2/df	247.2/98 = 2.52
	CFI	0.894

	RMSEA (95% CI)	0.099 (0.084-0.114)
	SRMR	0.079

Table 9. SIMS goodness-of-fit indices from CFA models of the *short-term* context data.

SIMS Short-term Context		
Model Fit Indices	BIC	8025.6
	χ^2/df	146.2/98 = 1.49
	CFI	0.954
	RMSEA (95% CI)	0.056 (0.036-0.074)
	SRMR	0.082

3.5. Construct Validity – Relationship to Other Variables Evidence

MCMS and AMS are grounded in the same theoretical framework as SIMS (self-determination theory) and also seek to measure IM, IR, ER, and AM, which facilitated direct comparisons between these instruments and SIMS. MBI-SSEE seeks to measure emotional exhaustion, which is similar albeit distinct from amotivation.

The relationship of SIMS variables to those of AMC, MCMS, and MBI-SSEE were assessed by examining simplex patterns of internal Pearson's r correlations and reliability as reflected by Cronbach's alpha.

In the post-situational context, Cronbach's alpha values were 0.93 (IM), 0.98 (IR), 0.89 (ER), and 0.98 (AM) for MCMS; 0.96 (IM), 0.96 (IR), 0.95 (ER), and 0.31 (AM) for AMS; and 0.98 (AM) for MBI-SSEE (Table 10). In the short-term context, Cronbach's alpha values were 0.98 (IM), 0.96 (IR), 0.88 (ER), and 0.99 (AM) for MCMS; 0.91 (IM), 0.97 (IR), 0.96 (ER), and 0.84 (AM) for AMS; and 0.93 (AM) for MBI-SSEE (Table 11). Indicating that the internal consistency of MCMS, AMS, and MBI-SSEE was acceptable given that all but one of the Cronbach's alpha values exceeds the ≥ 0.7 minimum criterion [26–29]. The notable exception being for the AM dimension of the AMS scale (Cronbach's alpha = 0.31).

Table 10. MCMS, AMS, and MBI-SSEE Cronbach's alpha values and Pearson's r correlation values relative to SIMS calculated from the *post-situational* context data.

Post-situational	SIMS	IM	IR	ER	AM	Cronbach's alpha
MCMS	IM	0.85	0.65	-0.21	-0.29	0.93
	IR	0.17	0.23	0.06	-0.08	0.98
	ER	-0.09	-0.01	0.30	0.23	0.89
	AM	-0.40	-0.55	0.32	0.76	0.98
AMS	IM	0.64	0.58	-0.21	-0.16	0.96
	IR	0.44	0.51	-0.11	-0.20	0.96
	ER	0.14	0.08	0.10	0.21	0.95
	AM	-0.34	-0.49	0.29	0.81	0.31
MBI-SSEE		-0.40	-0.33	0.23	0.37	0.998

Table 11. MCMS, AMS, and MBI-SSEE Cronbach's alpha values and Pearson's r correlation values relative to SIMS calculated from the *short-term* context data.

Short-term	SIMS	IM	IR	ER	AM	Cronbach's alpha
MCMS	IM	0.86	0.30	-0.22	-0.18	0.98
	IR	-0.09	0.30	0.24	-0.15	0.96
	ER	-0.07	0.01	0.32	0.05	0.88
	AM	-0.32	-0.29	0.21	0.59	0.99

AMS	IM	0.64	0.31	-0.12	0.02	0.91
	IR	0.18	0.39	0.25	-0.18	0.97
	ER	-0.14	0.11	0.35	0.12	0.96
	AM	-0.1	-0.31	0.13	0.65	0.84
MBI-SSEE		-0.45	-0.10	0.23	0.31	0.93

The Pearson's r correlations between SIMS and the MCMS and AMS scales demonstrated simplex trends in both post-situational and short-term contexts. Dimensions of the motivation construct that were theoretically close (e.g., like IM and IR) had positive correlations, whereas those that were farther apart (e.g., like IM and AM) exhibited negative correlations (Tables 10 and 11). Comparison of SIMS to the MBI-SSEE likewise revealed a pattern of correlation, with IM and IR showing negative correlation with emotional exhaustion, and ER and AM demonstrating a positive correlation (Tables 10 and 11).

3.6. Construct Validity – Consequences Evidence

No direct, negative repercussions of the SIMS instrument being used for its intended purpose in a medical school context were identified by the authors (B.W. and D.M.), the focus group participants, or the subject matter experts. Student participants understood through informed consent that their participation was entirely anonymous and optional, and that participation or lack thereof would have no impact on their progression through and completion of medical school.

4. Discussion

Greater emphasis is being placed on the interplay between student motivation and educational methods across a range of health professions, including medicine [35–40]. For example, motivational pedagogy refers to teaching methods and educational practices that aim to enhance students' intrinsic motivation to learn. This method is based on the conviction that motivation is an essential component of the learning process, and educators can profoundly influence student engagement and success by nurturing an educational atmosphere that encourages motivational growth [3,41]. As an educational strategy, it necessitates a reliable means to measure and track student motivation to ensure that teaching methods are effectively tailored to enhance engagement and learning, allowing for iterative refinements that address individual motivational needs, promote optimal educational outcomes, and reduce the risk of medical student burnout. SIMS is an instrument suited to the measurement of motivation within the context for which it was designed [15]. However, when survey instruments such as SIMS are moved to different contexts, the underlying assumptions, population characteristics, validity evidence, and relevance of the construct it was originally designed to measure may no longer hold true. Accordingly, this study sought validity evidence for the use of the SIMS in accurately measuring learner motivation among medical students in pre-clerkship short-term and post-situational contexts.

In this investigation, we evaluated the validity evidence for applying SIMS to gauge medical student motivation across two distinct contexts. The first context, referred to as '*post-situational*', adapted the survey's intended real-time assessment to a retrospectively proximal one, enabling the gathering of information on a student's motivation after the occurrence of a learning event, rather than during. This approach provides a more feasible alternative for collecting motivational data; it minimizes disruption to the learning experience and allows for more reflective and accurate self-assessment by the respondents of their motivational states, free from the immediate pressures or engagements of the learning activity. The second context, designated as '*short-term*', assessed a student's motivation to study over a one-week period. Medical students must engage in self-directed studying after learning events in order to consolidate their understanding, challenge their mastery of the material, and bridge the gap between theoretical knowledge and practical application, a crucial step in developing the competence and confidence required for clinical practice [42,43].

Expecting perfect validity evidence for a survey instrument is impractical due to several factors: validity is multi-dimensional; it is influenced by the variability in respondents' behaviors and the

characteristics of the sample used for validation. Additionally, no survey can entirely escape measurement error, and the statistical methods used to assess validity are themselves imperfect. Thus, while high-quality validity evidence is the goal, perfection in validity is an ideal rather than an attainable standard [18,44–46]. With this in mind, we have gathered a considerable body of construct validity evidence supporting the use of SIMS to measure motivation in pre-clerkship medical students in post-situational and short-term contexts. The *content*, *response process*, and *consequences* evidence were acceptable for both use case scenarios. For *relationship to other variables* evidence, this was also substantially true. The Pearson's r correlation values and simplex patterns were acceptable for both contexts, but the post-situational Cronbach's alpha for the measurement of amotivation (AM) with the AMS survey tool was 0.31 indicating low internal reliability of the measurement of this dimension of motivation with this particular instrument. This is more of a reflection on AMS in the two medical student contexts rather than SIMS. Lastly, for *internal structure* evidence, the Cronbach's alpha and Pearson's r correlational data and simplex patterns were acceptable for both use case scenarios; indicating good internal consistency and reliability. All CFA hypothesized factor loadings were found to be significant ($p < 0.01$) for both contexts. For the short-term context, the CFA model fit indices data were excellent. However, for the post-situational context, the CFA model fit indices data fell short of excellent. χ^2/df ($\chi^2/df = 2.52$) was higher than the ideal fit value threshold of < 2.0 , but was indicative of the model fitting reasonably well; the CFI value (CFI = 0.894) was also below the preferred threshold, suggesting the model may not fit as well as it could; the RMSEA value (RMSEA = 0.099) was also a bit higher than what we would want for a good fit; whereas the SRMS (SRMR = 0.079) was just at the cutoff of what is generally considered good, suggesting an adequate model fit. In summary, the post-situational CFA data represented a model that had an acceptable fit in some respects (i.e., factor loadings and SRMS) but indicated it might be worth examining the model to see if it could be improved (i.e., χ^2/df , CFI, RMSEA). This was consistent with the MI data for the post-situational context that proposed that one IM question be re-assigned to the IR cohort of questions, and another IR question might cross-load onto the IM dimension.

Hence, we can deduce that the SIMS instrument is effective in differentiating among the four dimensions of motivation delineated by SDT when administered to pre-clerkship medical students in a short-term context. The instrument's application in a post-situational context should be approached more cautiously due to its limited capacity to discriminate between identified regulation (IR) and intrinsic motivation (IM). However, that said, a less than ideal CFA model fit in distinguishing between IR and IM within the framework of SDT is not particularly concerning, given their conceptual adjacency and that both represent self-determined forms of motivation [4].

4.1. Limitations and Threats to Validity

This study was conducted at a single U.S. medical school, which may limit its generalizability to other domestic and international medical institutions. Nonetheless, the underlying concepts are widely relevant, and tools based on self-determination theory, such as SIMS, have been utilized in diverse research contexts worldwide. Potential validity threats like the Hawthorne effect were mitigated by the fact that students participated in a familiar environment and were accustomed to completing surveys as a part of their routine activities. Furthermore, to reduce response bias that could skew the psychometric survey outcomes, participants were not informed of the anticipated results, and the surveys were conducted anonymously.

5. Conclusions

While it is commonly believed that medical students possess a high level of inherent motivation, even a cursory observation indicates that their motivation, much like that of any other students, fluctuates over time. Gaining a deeper understanding of this dynamic and identifying effective intervention strategies could greatly benefit medical educators in a variety of areas, from curriculum development to lecture design. This investigation adds to the extensive research on self-determination theory in education and enriches the more limited body of work specific to medical education. By evaluating the application of SIMS in two novel contexts, this research not only bolsters

the existing validation evidence for SIMS but also enhances its practical applicability in medical education. Considering the results of this study, educators can use SIMS with increased assurance to explore the motivational dynamics of pre-clerkship medical students in short-term academic scenarios.

Author Contributions: Conceptualization, D.M. and B.W.; data curation, D.M. and B.W.; formal analysis, D.M. and B.W.; funding acquisition, D.M. and B.W.; investigation, D.M. and B.W.; methodology, D.M. and B.W.; project administration, D.M.; resources, D.M. and B.W.; software, D.M. and B.W.; supervision, D.M.; visualization, D.M. and B.W.; writing—original draft, D.M. and B.W.; writing—review and editing, D.M. and D.M. All authors have read and agreed to the published version of this manuscript.

Funding: This research was funded by the internal ‘Scholarly Reflection and Concentration/Capstone (SRCC)’ capstone project budget of the Frank H. Netter MD School of Medicine at Quinnipiac University.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Quinnipiac University (#06617).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available in order to maintain participant confidentiality per IRB protocol.

Acknowledgments: The authors wish to acknowledge Dr. Richard Feinn (Professor of Medical Sciences, Frank H. Netter MD School of Medicine at Quinnipiac University), as our consulting statistician, for his guidance with the statistical analyses presented in this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This is a reproduction of the original version of SIMS published by Guay et al. (2000) [15].

Appendix B

This is an example of the four survey instruments used to gather data for the *post-situational* context: SIMS, MCMS, AMS, and MBI-SSEE.

Appendix C

This is an example of the four survey instruments used to gather data for the *short-term* context: SIMS, MCMS, AMS, and MBI-SSEE.

Appendix D

These are supplemental figures generated by SPSS Statistics, Version 27 for the *post-situational* and *short-term* context CFA.

References

1. Oxford English and Spanish Dictionary. “motivation, n.” (2021). Retrieved November 1, 2023, from <https://www.lexico.com/en/definition/motivation>
2. Deci, E., & Ryan, R. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. Springer.
3. Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67. <https://doi.org/10.1006/ceps.1999.1020>
4. Deci, E. L., & Ryan, R. M. (2000). The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
5. Mayer, R. E. (2008). *Learning and instruction*. Prentice Hall.
6. Dyrbye, L. N., Thomas, M. R., & Shanafelt, T. D. (2006). Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. *Academic Medicine*, 81(4), 354–373. <https://doi.org/10.1097/00001888-200604000-00009>
7. Brazeau, C., Shanafelt, T. D., Durning, S. J., Massie, F. S., Eacker, A., Moutier, C., Satele, D., Sloan, J. A., & Dyrbye, L. N. (2014). Distress among matriculating medical students relative to the general population. *Academic Medicine*, 89(11), 1520–1525. <https://doi.org/10.1097/acm.0000000000000482>

8. Mata, D. A., Ramos, M. A., Bansal, N., Khan, R., Guille, C., Di Angelantonio, E., & Sen, S. (2015). Prevalence of depression and depressive symptoms among resident physicians. *JAMA*, *314*(22), 2373. <https://doi.org/10.1001/jama.2015.15845>
9. Cooke, M., Irby, D. M., Sullivan, W., & Ludmerer, K. M. (2006). American medical education 100 years after the Flexner report. *The New England journal of medicine*, *355*(13), 1339–1344. <https://doi.org/10.1056/NEJMra055445>
10. Drolet, B. C., & Rodgers, S. (2010). A Comprehensive Medical Student Wellness Program—Design and Implementation at Vanderbilt School of Medicine. *Academic Medicine*, *85*(1), 103–110. <https://doi.org/10.1097/acm.0b013e3181c46963>
11. Slavin, S., Schindler, D., & Chibnall, J. T. (2014). Medical Student Mental Health 3.0. *Academic Medicine*, *89*(4), 573–577. <https://doi.org/10.1097/acm.0000000000000166>
12. Dyrbye, L. N., West, C. P., Satele, D., Boone, S., Tan, L., Sloan, J. A., & Shanafelt, T. D. (2014). Burnout among U.S. medical students, residents, and early career physicians relative to the general U.S. population. *Academic Medicine*, *89*(3), 443–451. <https://doi.org/10.1097/acm.0000000000000134>
13. Binning, J. (2016). “Construct”. Encyclopedia Britannica. Retrieved November 1, 2023, from <https://www.britannica.com/science/construct>
14. DeVellis, R. F. (2016). *Scale development: Theory and Applications*. SAGE Publications.
15. Guay, F., Vallerand, R., & Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*(3), 175–213. <https://doi.org/10.1023/A:1005614228250>
16. Jenö, L. M., Raaheim, A., Kristensen, S. M., Kristensen, K. D. B., Hole, T. N., Haugland, M., & Mæland, S. (2017). The Relative Effect of Team-Based Learning on Motivation and Learning: A Self-Determination Theory Perspective. *CBE- Life Sciences Education*, *16*(4), ar59. <https://doi.org/10.1187/cbe.17-03-0055>
17. Zimmerman, B. J. (2008). Investigating Self-Regulation and Motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. <https://doi.org/10.3102/0002831207312909>
18. Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education.
19. Downing, S. M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, *37*(9), 830–837. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>
20. Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and application. *The American Journal of Medicine*, *119*(2), 166.e7-166.e16. <https://doi.org/10.1016/j.amjmed.2005.10.036>
21. Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: *Measurement and Evaluation in Counseling and Development*, *36*(3), 181–191. <https://doi.org/10.1080/07481756.2003.11909741>
22. Vallerand, R. J., & Fortier, M. N. (1998). Measures of Intrinsic and Extrinsic Motivation in Sport and Physical Activity: A Review and Critique. In J. L. Duda (Ed.), *Advances in Sport and Exercise Psychology Measurement* (pp. 81-101). Morgantown, WV: Fitness Information Technology.
23. Schaufeli, W. B., Martínez, I., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002). Burnout and engagement in university students. *Journal of Cross-Cultural Psychology*, *33*(5), 464–481. <https://doi.org/10.1177/0022022102033005003>
24. Posch, L., Bleier, A., Lechner, C. M., Danner, D., Flöck, F., & Strohmaier, M. (2017). Measuring Motivations of Crowdworkers: The Multidimensional Crowdworker Motivation Scale. *arXiv (Cornell University)*. <http://export.arxiv.org/pdf/1702.01661>
25. Muthen, L. K., & Muthen, B. O. (2017). *MPlus Version 8 User's Guide*.
26. Nunnally, J., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill Humanities/Social Sciences/Languages.
27. Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *BMJ*, *314*(7080), 572. <https://doi.org/10.1136/bmj.314.7080.572>
28. DeVellis, R. F. (2003). *Scale development: Theory and Applications*. Sage Publications.
29. *Using and interpreting Cronbach's Alpha* | UVA Library. (n.d.). Retrieved November 1, 2023 from <https://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/>
30. Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111–150. https://doi.org/10.1207/s15327906mbr3601_05
31. Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, *6*(1). <https://doi.org/10.21427/d7cf7r>
32. Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*, Fourth edition. Guilford Publications.

33. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
34. *Modification indices | Just Enough R.* (n.d.-b). Retrieved November 1, 2023 from <https://benwhalley.github.io/just-enough-r/modification-indices.html#:~:text=T%20he%20modification%20index%20is,was%20added%20or%20constraint%20free d.>
35. Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, 26(3-4), 325-346. doi:10.1080/00461520.1991.9653137
36. Ten Cate, O., Snell, L., & Carraccio, C. (2010). Medical education: Past, present and future. *Perspectives on Medical Education*, 29(6), 320-327. doi:10.1007/s40037-010-0036-9
37. Kusrkar, R. A., Croiset, G., & Cate, O. T. (2011). Twelve tips to stimulate intrinsic motivation in students through autonomy-supportive classroom teaching derived from Self-Determination Theory. *Medical Teacher*, 33(12), 978–982. <https://doi.org/10.3109/0142159x.2011.599896>
38. Al-Eraky, M. M., & Marei, H. F. (2016). A fresh look at Miller’s pyramid: Assessment at the “Is” levels. *Medical Teacher*, 38(8), 830-836. doi:10.3109/0142159X.2015.1112894
39. Mann, K. (2004). The role of educational theory in continuing medical education: Has it helped us? *Journal of Continuing Education in the Health Professions*, 24(Supplement 1), S22–S30. <https://doi.org/10.1002/chp.1340240505>
40. Kusrkar, R., Ten Cate, T., Van Asperen, M., & Croiset, G. (2011). Motivation as an independent and a dependent variable in medical education: a review of the literature. *Medical Teacher*, 33, 242–262. <https://doi:10.3109/0142159X.2011.558539>
41. Stipek, D. J. (2002). *Motivation to learn: Integrating Theory and Practice.* Allyn & Bacon.
42. Blumberg, P. (2000). Evaluating the Evidence That Problem-Based Learners Are Self-Directed Learners: A Review of the Literature. In D. H. Evensen & C. E. Hmelo (Eds.), *Problem-Based Learning: A Research Perspective on Learning Interactions* (pp. 199-226). Lawrence Erlbaum Associates.
43. Murad, M. H., & Varkey, P. (2008). Self-directed learning in health professions education. *Annals of the Academy of Medicine, Singapore*, 37(7), 580.
44. Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <https://doi.org/10.1037/h0040957>
45. Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73. <https://doi.org/10.1111/jedm.12000>
46. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.