



## OPEN ACCESS

## EDITED BY

Douglas F. Kauffman,  
Medical University of the Americas – Nevis,  
United States

## REVIEWED BY

Abdolreza Khalili,  
Urmia University, Iran  
Abdollah Bapiri,  
Urmia University, Iran

## \*CORRESPONDENCE

Oun Almesaar  
✉ [omesaar@taibahu.edu.sa](mailto:omesaar@taibahu.edu.sa)

RECEIVED 26 June 2023

ACCEPTED 25 January 2024

PUBLISHED 07 February 2024

## CITATION

Almesaar O (2024) L2MSS in the Saudi Arabian  
context: item, scale, and external validity  
analysis.

*Front. Educ.* 9:1247701.

doi: 10.3389/feduc.2024.1247701

## COPYRIGHT

© 2024 Almesaar. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# L2MSS in the Saudi Arabian context: item, scale, and external validity analysis

Oun Almesaar \*

Taibah University, Madinah, Saudi Arabia

With English being the modern *lingua franca*, educational policies in Saudi Arabia were recently changed in order to provide better English language education to Saudi students. In spite of the government's efforts to improve English proficiency, most Saudi students do not benefit sufficiently from the new curricula. Language 2 Learning Motivation Self-System (L2MSS), proposed by Dörnyei, provides a potential framework for understanding determinants of English proficiency among Saudi students. The pilot study first addressed the internal reliability, factor structure and item-level performance (classical model and IRT) of two new questionnaires developed to complement L2MSS with measures of cognitive, emotional, and behavioral aspects of attitude towards English. The main study focused on relationship between L2MSS and related variables with performance on an English test based on TOEIC. The study has shown that, while Saudi students tend to be motivated to learn English, L2MSS scores are not related to performance on an objective test of English proficiency. Moreover, the two new questionnaires, while associated with L2MSS, are not associated with English test performance. Exploratory factor analysis (with maximum likelihood extraction and direct oblimin rotation) of all L2MSS items resulted in a general L2MSS factor (29% variance), factors 2 (7% variance) and 3 (3%) had mixed loadings of "ought-to-self" and L2 experience items, while factor 4 (2.7%) mainly had unique ideal self-loadings.

## KEYWORDS

L2MSS, motivation, English language, Saudi Arabia, education

## 1 Introduction

Throughout the second half of the 20th century, English language was slowly making its way into Saudi Arabian school curricula and society at large. In the early 2000s, English language classes were introduced to all primary schools at grade 6 (Khawaji, 2022, p. 267). Vision 2030, a grand plan for social, educational, and economic reform in Saudi Arabia, more specifically its Human Capability Development Program, foresees the introduction of English language classes from grade 1. Vision 2030 is a logical conclusion of the gradual opening up of the Saudi Arabia to the world, spurred by the discovery of major oil reserves and the exponential economic development that ensued. In such a context, the English language became not simply an instrument of cultural exchange, but also an indispensable tool of economic exchange, high English competence becoming a vital asset within a multitude of modern workplaces (Khawaji, 2022, p. 266).

Saudi Arabia also has a thriving private education sector (Faruk, 2013, p. 76; Al-Seghayer, 2014, p. 17), with numerous children receiving private English classes. For the time being, English is still the only foreign language taught in public schools in Saudi Arabia, a sign of the privileged status of the English language in the country.

However, it seems that the overall English competence of Saudi students remains fairly low (Hamouda, 2013, p. 17; Alrabai, 2016, p. 21; Alshammari, 2022, p. 129). Considering the global importance of English as a *lingua franca*, as well as the decades of initiatives to improve English competence among Saudi students, we aim to identify some of the possible factors contributing to the explanation of low English proficiency among Saudi students, focusing on L2 Motivational Self-System-L2MSS (Dörnyei, 2009). Identifying the factors behind low English competence is not only a task of scientific importance, but also something that can inform new educational reforms and maximize the outcomes of educational investments.

## 1.1 L2MSS

The introduction of L2MSS came as an adaptation of Gardner's integrative orientation ("integrativeness"), which has been used previously to explain people's motives to learn a foreign language (Gardner, 2001). Gardner's theory (Gardner, 2001) has mainly been tested in the Canadian context, where people are motivated to learn another language (English or French) so that they can communicate with their fellow citizens. Dörnyei argued that, in most countries, students are learning a foreign language that is not used as a native language by a significant part of the domestic population (Dörnyei, 2009, p. 24). As a Hungarian, Dörnyei was well aware that most students around the world acquire a foreign language (English most often) which is not used for everyday communication in their native countries. The need to integrate to a group, therefore, does not play an important part in most cases of learning a foreign language. Dörnyei's work (Dörnyei, 2009, 2019) can therefore be understood as a set of attempts to expand the concept of L2 motivation to a broader, more global context.

Dörnyei (2009) developed a new approach to systematizing second language learning motivation, which relies heavily on the concept of self. This approach has since been referred to as L2MSS. Dörnyei (2009, 2019) built L2MSS on previous studies of L2 acquisition, introducing concepts like "possible selves" and applying self-discrepancy theory to the study of L2 learning motivation.

Crucially, he introduced the notions of "ideal L2 self" and "ought-to L2 self," which are related to people's motivation to learn a second language. Ideal L2 self is a component of a broader ideal self which consists of representations of success, prosperity, and happiness that one wishes to attain (Dörnyei, 2009, p. 29). Thus, ideal L2 self relates to a person's vision regarding the mastery of a foreign language. Ideal L2 self, as is the case with ideal self in general, relates to individual wishes and aspirations. On the other hand, ought-to L2 self consists of others' expectations regarding a person's foreign language competence. In other words, ought-to self is what we think others want us to be (Dörnyei, 2009, p. 29).

Both ideal and ought-to self are internalized, meaning that we consider them to be parts of our personalities. Moreover, the two selves affect each other; for instance, what others expect from us can gradually become a component of our ideal self and vice versa.

Dörnyei (2009, 2019) introduced another important factor – L2 Learning Experience (L2LE). While the selves are related to future and involve imagery, L2LE refers to presence of prerequisites for learning L2, such as teaching materials, peers, teachers, etc. (Dörnyei, 2019, p. 25). Motivation depends not only on a person's representations of

themselves (ought-to and ideal selves) but also on their experience with L2, which depends very much on the learning materials, classmates, etc.

These three main components (ideal L2 self, ought-to L2 self, and L2LE) relate to a learner's intended learning efforts (ILEs), with ILEs often being used as a criterion measure for the external validity of L2MSS (Moskovsky et al., 2016, p. 643). ILEs in the sphere of L2 motivation are exemplified by the following statements: "If they offer additional night-time classes I would like to take them," while actual learning efforts relate to statements like "I am taking additional night-time classes." (Moskovsky et al., 2016, p. 643).

L2MSS is likely associated with numerous well-established constructs coming from the broad field of research on L2 acquisition. For instance, it has been shown that components of L2MSS can predict L2 willingness to communicate (WTC) (Amirian et al., 2021, p. 12; Li and Liu, 2021; Ebn-Abbasi et al., 2022; Zhou, 2022). Furthermore, ideal self is associated with growth mindset, a concept used to explain academic engagement (Sadoughi et al., 2023). Connections between L2MSS and other concepts were also explored, such as academic resilience and more specifically the factor of perseverance (Çelen, 2020). A recent study has also shown that there are significant connections between L2MSS and L2 anxiety (Sadoughi and Hejazi, 2023). There have also been initiatives of reconceptualization of L2MSS, chiefly materialized in the introduction of another component, "feared L2 self" (Peker, 2020).

Boo et al. (2015) concluded that L2MSS is the dominant approach in the field. Al-Hoorie (2018) conducted the first meta-analysis of L2MSS studies – analyzing the results of 32 research reports, with a sample of 32,072 language learners. It has been found that ideal L2 self, ought-to L2 self, and L2LE are all positively and strongly associated with ILEs ( $r$  coefficients vary from 0.379 to 0.656). However, these variables mostly have rather weak relationships with arguably the most important criterion variable – language achievement (measured via academic achievement or performance on an objective test); even the relationship between ILEs and language achievement is weak and non-significant (Al-Hoorie, 2018). It has to be emphasized that ILEs are regarded as "closest" to actual behavior by Dörnyei (2009) and are supposed to mediate the relationships between other components of L2MSS and actual behavior (e.g., achievement). However, it appears that ILEs, alongside other components of L2MSS, are unrelated to L2 achievement and performance.

An important point to note is that in Al-Hoorie's meta analysis (Al-Hoorie, 2018), only two out of 32 reviewed studies included administration of an objective English test as a validation of L2MSS. Lamb (2012) administered the C-test (a form of word-completion task) developed for the purposes of the study, and Moskovsky et al. (2016) administered a test (reading and writing), based on IELTS. A portion of Lake's participants (Lake, 2013) reported on their TOEIC-Bridge test scores, but this author did not directly assess participants' L2 proficiency. In other words, there is a paucity of studies juxtaposing L2MSS and L2 proficiency measured via an objective L2 test.

Moskovsky et al. (2016) tested the relevance of L2MSS in the Saudi context by introducing a crucial measure of language performance – an objective English test. Their conclusions were similar to those of Al-Hoorie (2018); namely, L2MSS seems to be unrelated to English language performance. Moskovsky et al. (2016) administered a short English test, which was based on IELTS,

focusing on reading and writing abilities. These authors analyzed zero-order correlations between L2MSS components and performance on the English test, finding no significant relationships between L2MSS components and English proficiency. The study by [Moskovsky et al. \(2016\)](#) is important because it points to an important way of validating L2MSS by testing its relationship with performance on an objective language test. This way, L2 proficiency can be directly assessed instead of relying on academic achievement. Additionally, [Moskovsky et al. \(2016\)](#) drew their test from the IELTS workbook; the reading subtask involved the reading of a 700-word text and responding to 15 questions relating to the text; the writing subtask involved writing a 200-word text on a neutral topic.

[Lamb \(2012\)](#) developed a somewhat different method of assessing English proficiency. In this study, the so-called C-test was used, which consisted of five short texts. Within each text, roughly 50% of words were incomplete, and the task was to fill in the missing letters.

It is true that administering objective L2 tests is time-demanding and rather challenging for participants, which is perhaps one of the reasons why in most studies identified by [Al-Hoorie \(2018\)](#), authors mostly turned toward academic achievement for validating L2MSS. This usually means obtaining self-report measures of academic achievement, which is potentially problematic as students might be tempted to present themselves in a more positive light (in case they have bad grades). Moreover, it is possible that there are problems with the English proficiency assessment system in Saudi Arabia ([Almuhaimed, 2002](#)). For these reasons, it is crucial to test the validity of L2MSS by administering an objective English test to participants akin to those used by [Lamb \(2012\)](#) and [Moskovsky et al. \(2016\)](#).

Besides concerns about the actual relevance of [Dörnyei's](#) approach, there are also potential concerns about the factor structure of the scales developed to assess [Dörnyei's](#) constructs ([Dörnyei, 2009, 2019](#)). It has to be noted that factor analyses of L2MSS are somewhat scarce ([Csizér, 2019](#)) and that there are few studies examining the interrelationships between the L2MSS components. An exploratory factor analysis of ideal and ought-to scales yielded two main factors that, indeed, corresponded to the two “self” constructs ([Thompson and Erdil-Moody, 2016](#)), although the researchers did not include items relating to L2LE and ILEs. [Kim and Kim \(2021\)](#) were also able to identify ideal and ought-to latent factors, although they did not include L2LE, nor ILEs. [Kang \(2017\)](#) conducted exploratory factor analysis of a Korean version of the L2MSS scale, identifying 8 factors. Most recently, [Lee and Lee \(2021\)](#) were able to identify latent factors corresponding to ideal and ought-to selves. We have to mention that these authors used different operationalizations of L2MSS, and this can be an important source of variance in the factor structure of L2MSS between studies.

## 1.2 Attitude towards English and action plans

Attitude can be defined as one's evaluation of a stimulus object ([Ajzen, 2005](#)). The importance of attitudes in foreign language learning was noted a long time ago ([Smith, 1971](#)). [Smith \(1971\)](#) gave an explanation of the attitude towards learning a foreign language which encompassed cognitions, feelings, and behavioral intentions. The tripartite view of attitudes (including cognitive, affective, and

behavioral components) gained prominence a long time ago ([Ostrom, 1969; Bagozzi et al., 1979](#)) and continues to hold relevance in more recent times ([Kaiser and Wilson, 2019](#)) in a variety of different fields ([Fatimah, 2015; Kim et al., 2015](#)). The tripartite attitude model is also an important way to systematize attitudes towards learning a foreign language ([Tseng, 2010; Lasagabaster, 2017](#)). [Tseng \(2010\)](#), for instance, found that the tripartite model of attitudes toward English vocabulary learning has greater explanatory power compared to a one-factor model.

Considering that the L2MSS questionnaire already has a fair number of items relating to behavioral intentions (Intended efforts subscale), we will attempt to construct a questionnaire of attitude toward English language by focusing on cognitive and emotional aspects of this attitude.

[Dörnyei \(2009\)](#) suggested that very concrete Action Plans may be important for bridging the gap between L2MSS components and relevant behavior. If concrete Action Plans are related to actual behavior, they can possibly act as mediators between components of L2MSS and actual behavior. [He et al. \(2011\)](#) examined students' strategies relating to writing in English. These authors used the think-aloud protocols to elicit students' main writing strategies, initially identifying 21 specific writing strategies. Upon further analysis these specific strategies were classified into five main groups: planning, monitoring and evaluating, revising, retrieving, and compensating. ‘Organizing thoughts’ and ‘thinking about the intended audience’ are examples of sub-strategies under planning. ‘Monitoring and evaluating’ refers to actions such as self-monitoring of the use of various expressions, self-monitoring of the organization of ideas and plans, or eliciting questions about the quality of writing. Revising grammar, punctuation, spelling, structure, are examples of the broader ‘revising’ strategy. ‘Retrieving’ refers to retrieving expressions from one's vocabulary and comparing them with the intended expressions, as well as retrieving background knowledge necessary for constructing the overall structure and rationale of the text. Finally, examples of ‘compensating’ strategies are consulting dictionaries, translating intended expressions into English, and finding words with similar meanings when the target word cannot be retrieved. Students' Action Plans can be systematized with the help of these five main strategies, and one of the aims of this study is to construct and test the Action Plans questionnaire based on the research of [He et al. \(2011\)](#).

## 1.3 Research gap and contribution of the study

It can be stated generally that authors who researched L2MSS mostly focused on the three main components (ideal L2 self, ought-to L2 self, and L2LE) and how they relate to ILEs and other conceptually close constructs, such as WTC. The value of the whole model, however, relies on the link with actual behavior (achieving L2 proficiency). As we have seen, there are indications that L2MSS is not related to performance on English tests ([Moskovsky et al., 2016; Al-Hoorie, 2018](#)). However, studies that attempt to link L2MSS with English test performance, as we have also seen, are scarce, with most authors focusing on academic grades ([Al-Hoorie, 2018](#)). This is why further L2MSS studies need to include L2 proficiency variables, measured via valid and widely accepted language tests.

[Thompson and Erdil-Moody \(2016\)](#) emphasized the need to perform exploratory factor analyses of L2MSS scales in different

cultural contexts. While there are studies employing L2MSS in the Saudi context (Eusafzai, 2013; Moskovsky et al., 2016; Alqahtani, 2020), not all of them explored the factor structure of L2MSS. Furthermore, Moskovsky et al. (2016) were only interested in the factor structure of two subscales (L2LE and ILEs). Eusafzai (2013) was able to identify 8 distinct components of the L2MSS scale, utilizing Principal Component Analysis (PCA), which is arguably not the ideal approach for identifying the true latent structure of a dataset; PCA is mainly a method of reducing a set of highly correlated variables to a smaller set (Suhr, 2005), while true factor analysis employs various extraction methods (e.g., maximum likelihood) and rotations to identify the underlying latent structure of a dataset. It can be concluded that a comprehensive analysis of the factor structure of L2MSS scales in the Saudi context is yet to be carried out.

Moskovsky et al. (2016, p. 11), who tested L2MSS in Saudi Arabia, suggested that “those with high proficiency may not perceive English language learning as effortful behavior, but rather as a pleasurable and fun activity.” In other words, it is possible that L2MSS is not able to account for this dimension of relationship towards a foreign language, which can be described as attitude towards a foreign language. L2MSS is a rather comprehensive questionnaire with numerous items, but only a few of them cover the attitude towards a foreign language (e.g., “I really enjoy learning English,” which is an item of the L2LE subscale). Attempting to account better for this dimension of relationship toward a foreign language could potentially increase the explanatory power of L2MSS model.

Finally, we have mentioned Dörnyei's (2009) suggestion regarding the addition of concrete Action Plans to L2MSS. However, L2MSS studies scarcely addressed this potentially important concept. It may seem that intended learning efforts (“I am working hard at learning English” or “I think that I am doing my best to learn English”), as operationalized by Moskovsky et al. (2016), nicely cover students' Action Plans, but this is not the case. The ILE items are fairly broad and only address very general intentions, not concrete action plans. Thus, a new Action Plans scale needs to be constructed in order to test the contribution of this concept to L2MSS.

## 1.4 Research aims and hypotheses

The purpose of this study is to test the performance of the L2MSS operationalization of L2MSS (Moskovsky et al., 2016), more specifically its factor structure, and its criterion validity in the Saudi context (relationship with performance on an objective English test), and the performance of two new scales (“Attitude towards English” and “Action Plans”) and their relationships with L2MSS and objective performance. It is possible that the introduction of the two new constructs, which were previously left out of L2MSS studies, might provide the link between the main L2MSS constructs and real-life performance.

More specifically, it is hypothesized that:

1. Neither L2MSS components nor Intended efforts explain a significant portion of the variance of performance on an objective English test. [Based on the findings of Al-Hoorie (2018) and Moskovsky et al. (2016)].
2. Action Plans and Attitude towards English are associated with achievement on an objective English test.

3. Action Plans and Attitude towards English correlate highly with all L2MSS components.
4. Based on factor analyses of different operationalizations of L2MSS constructs, two main factors can be extracted, corresponding to ideal self and ought-to self (Thompson and Erdil-Moody, 2016; Kang, 2017; Kim and Kim, 2021; Lee and Lee, 2021).

## 2 Pilot study

A pilot study was conducted to test the reliability, factor structure, and item performance of two new questionnaires.

The initial phase of constructing the two new questionnaires involved the formation of two pools of items from which the final sets of items can be drawn. The researcher formed the two initial pools of items. A total of 33 items were formulated that potentially indicated cognitive or emotional components of the Attitude towards English language. The items referred to English grammar (7), pronunciation (7), phonology (7), orthography (7), as well as comparisons between English and other languages (5). Furthermore, 29 potential items that covered five strategies identified by He et al. (2011) were formulated. Thus we had items referring to planning (6), monitoring and evaluating (6), revising (6), retrieving (6), and compensating (5). Both pools of items were formulated having L2MSS items in mind, so that the content of new items does not overlap with the content of L2MSS, the goal of the two new questionnaires being to complement L2MSS.

These items were then presented to a brainstorming group (led by the researcher) involving 7 students who were all learning English as a second language. The students were asked about the relevance of the items for their relationship towards English language, and were asked to adjust the items as they saw fit; the comments and adjustments were considered by the reviewer, resulting in a total of 42 items for both scales. It was decided that students' input could potentially help us improve items and the face validity of the scales, because the primary goal of this study - and arguably the primary goal of the whole field of L2 research - is to aid the learning of a second language among students; thus the items should be recognized as relevant by the members of the target group.

The brainstorming group produced 22 items relating to Attitude towards English Language and 20 items relating to Action Plans. The researcher, with the help of two expert psychometricians, then considered these items and how they represent the underlying theoretical rationale of the new questionnaires. The most important question in this phase was whether the items reflected the underlying theoretical rationale upon which the two new questionnaires were built. The final version of the Attitude towards English scale had 10 items (5 for the cognitive and 5 for the affective component); the final version of the Action Plans questionnaire also had 10 items, 2 for each of the strategies mentioned above.

## 2.1 Method

### 2.1.1 Design

This is an exploratory study designed to test the scale-level and item-level performance of two new questionnaires.



### 2.1.2 Instruments

Attitude towards English Language consists of 10 items (5-point Likert scale) relating to cognitive and affective components of attitude towards English. The behavioral component of this attitude is already accounted for by the ILEs subscale of the L2MSS questionnaire as well as Action Plans.

The items were designed so that various aspects of cognitive and affective attitudes are covered. For instance, participants were asked about their attitude towards the English alphabet, the sound of the English language (e.g., “I like how the English language sounds”), and their general affective evaluation (e.g., “I like the English language”).

Participants also rated statements about the challenges related to the English language (e.g., “The English language is too complex for me”), as well as statements about the complexity of the English language and their curiosity to learn more (e.g., “The English language is very interesting”). There were three inverse items in total.

The Action Plans questionnaire tested the hypothesis that Action Plans may have an important role in materializing more abstract learning intentions and efforts. A new questionnaire was developed, based on He et al.'s (2011) strategies. Two items were dedicated to each of the five strategies: namely, planning (“I first have to think carefully about the topic at hand”); monitoring and evaluating (“I evaluate how the paragraph I am writing at the moment fits with the rest”); revising (“After I finish writing, I revise my text”); retrieving (“I try to remember how other people used a particular expression I intend to use”); and compensating (“When I am not sure about the spelling of a word, I try to find a synonym”).

### 2.1.3 Participants

It has been argued that, for proper one-parameter Item Response Theory (IRT) analysis of a questionnaire consisting of 10 items, 150 participants is enough to obtain stable estimates (Sahin and Anil, 2017), and this requirement was held in mind while gathering the sample.

A convenience pilot sample of 167 participants was gathered (Mean age = 19; 50% females). All participants studied at one of two Saudi Arabian universities and were native Arabic speakers. 30% studied Business Management and 70% studied Accounting. The English proficiency level was not assessed as part of the pilot study as the main aim was to assess the performance of the two new questionnaires.

### 2.1.4 Procedure

All participants completed questionnaires online in a single session via Survey Monkey. The data were obtained over a few days, in the final week of November 2021. Prior to completing the questionnaires, participants were required to read short instructions and sign an informed consent statement containing a short explanation of the purpose of the study and their basic rights as research participants. Participants were assured that their answers would remain anonymous.

### 2.1.5 Analyses

IRT analysis was performed via the Winsteps [program developed by Rasch (1960) and further elaborated by Linacre and Wright (2000)], providing new insights into how our questionnaires function. The Winsteps program only calculates Rasch's beta parameter, indicating a person's position on a latent variable (i.e., personality trait,

motivation, etc.). There are many ways for IRT analysis to inform the construction of questionnaires, but one of its biggest contributions is the estimation of infit and outfit for items and persons. A range between 0.7 and 1.3 is generally considered an interval of good fit (Linacre and Wright, 2000).

A classical item-level analysis was performed via MACRO RTT10G script for SPSS (Fajgelj and Kosanović, 2001). This item-level analysis included these metrics:

1. Keiser-Mayer-Olkin for representativeness (REP)
2. Squared Multiple Correlation for item reliability (REL)
3. Covariance of image variables (items) with their first principal component for homogeneity/validity (H)
4. Item-total correlation (Burt coefficient) for homogeneity/validity (B).

There are no standard guidelines for the interpretation of these metrics. A value of 0.3 is considered acceptable for item-level reliability, 0.5 for validity estimates (depending on the hypotheses about the number of underlying factors), and 0.9 is regarded as a satisfactory representativeness index (Fajgelj and Kosanović, 2001).

Descriptive statistical analyses were completed in SPSS (version 26), as well as exploratory factor analysis.

## 2.2 Results

A classical analysis of items was performed using RTT10G syntax for SPSS. Table 1 shows the main metrics for Action Plans items.

Items 2 (“I think about who is going to read my text before I start writing”), 4 (“I attempt to see whether my writing actually expresses my ideas”), and 10 (“When I am not sure about the spelling of a word, I try to find a synonym”) have very low reliability. The squared multiple correlations of these variables with other variables from the scale, are low. The validities—covariance between image items and first principal component (H)—and corrected item-total correlation (B) for these variables are also lower in comparison to other Action Plans items.

It was assumed that these issues arose due to the vagueness of items; because Item 10 contained two distinct statements, for example, participants might have found this confusing. Thus, these items were

TABLE 1 Item-level statistics for Action Plans items (classical model).

Item	REP	REL	H	B
Actp1	0.8956	0.3719	0.6813	0.6579
Actp2	0.8984	0.1089	0.4074	0.4552
Actp3	0.8916	0.3256	0.6440	0.6269
Actp4	0.8764	0.1574	0.4575	0.4870
Actp5	0.9080	0.3206	0.6485	0.6288
Actp6	0.9174	0.3143	0.6613	0.6382
Actp7	0.8699	0.3213	0.6276	0.6208
Actp8	0.8418	0.2601	0.5395	0.5518
Actp9	0.9041	0.2406	0.5865	0.5819
Actp10	0.8984	0.1700	0.5030	0.5262

rewritten—Item 2 “I think about the expectations and criteria of those who will read my text,” Item 4 “I compare all meanings of similar words I intend to use and think a lot about the best choice in the given circumstances,” and Item 10 “I think of all possible synonyms of a word before I make my choice.” The item-level analysis of attitude towards English items revealed no serious issues with item representativeness, reliability, and validity.

### 2.2.1 IRT analysis

The reliability metrics of Action Plans provided by Winsteps are: 0.66 for persons (separation = 1.41) and 0.89 for items (separation = 2.91). IRT item-level analysis of Action Plans revealed that items are too “easy”; item 5 is the worst in this respect (“After I finish writing, I revise my text”) and only accounted for the variation of the presence of Action Plans in two participants (with extremely low presence of this psychological construct). Most participants, even those with very low presence of Action Plans, agreed with this item. Items 1 (“I first have to think carefully about the topic at hand”), 6 (“I look carefully for grammar, spelling, and punctuation mistakes”), and 9 (“I consult English dictionaries to improve my word choices”) cover narrow spans of the Action Plans construct.

Item 5 was replaced with “I analyze each part of the text separately, trying to see if I can change the placement of certain paragraphs,” Item 1 with “First I write a very detailed plan consisting of numerous tasks and subtasks,” and Item 6 with “I spend a great deal of time looking for grammar, spelling, and punctuation mistakes.”

Three new items were added to the Action Plans scale to increase the diversity of items’ difficulties: “I send a text to my friends so that they can give me objective feedback”; “When I run into a problem I try to find help from native English speakers on the Internet”; and “I try to emulate the style of an English writer whose books I have read before.”

Regarding infit and outfit metrics, there are no issues with Action Plans items (all are between 0.7 and 1.3).

Person-level reliability for Attitude towards English was 0.81 (separation = 2.05), while item-level reliability was 0.97 (separation = 5.94). Items 1 (“I am curious about the logic behind the English language”) and 2 (“The English language is very interesting”) had infit and outfit outside the desired range. Item 4 (“Knowing English helps me learn new, fun things”) was easy. Thus, these items were replaced with: “I want to know everything about the logic behind the English language,” “The English grammar is much more interesting than other subjects like science or history,” and “I want to know all the little details of English grammar,” respectively.

### 2.2.2 Scale analyses

The mean for Attitude towards English was 38.5 (SD = 6.3), Action Plans had a mean of 42.2 (SD = 4.6), skewness and kurtosis were  $-0.645$  and  $0.344$ , for Attitude and  $-0.599$  and  $1.267$  for Action Plans. The Kolmogorov-Smirnoff test yielded a  $0.017$   $p$  value for Attitude and  $0.079$  for Action Plans; thus, Action Plans were normally distributed while Attitude towards English was not.

The questionnaires were analyzed by both classical and IRT methods. Reliability estimates under the classical model provided by SPSS (Cronbach’s alpha) were:  $0.866$  for attitudes towards English and  $0.769$  for Action Plans. The reliability of both questionnaires was acceptable, but item-level analyses helped us detect problematic items and further improve the questionnaires.

RTT10G syntax for SPSS also provided some general scale-level indicators, such as KMO representativeness. For the pilot version of Action Plans, KMO representativeness was  $0.897$  and attitude towards English was  $0.966$ . Thus, the choice of indicators for Action Plans is less representative of all possible items, something to note for subsequent scale adaptations.

Factor analysis of Action Plans scores was performed via maximum likelihood extraction, with direct oblimin rotation. The item sampling adequacy is solid (for Action Plans, KMO =  $0.897$ ), while Bartlett’s test of sphericity was significant. The analysis resulted in a one-factor solution, explaining  $27.5\%$  of variance (eigenvalue =  $2.752$ ).

The item sampling adequacy for Attitude towards English (KMO =  $0.966$ ) was very good, and Bartlett’s test of sphericity was significant. Attitude towards English was not normally distributed, meaning that maximum likelihood extraction cannot be used; therefore, an unweighted least squares extraction method was chosen alongside direct oblimin rotation. One factor was extracted, explaining  $43\%$  of variance (eigenvalue =  $4.304$ ).

The correlation between Attitude and Action Plans (final scores calculated as sums) was  $0.263$ .

## 2.3 Discussion

The pilot study was used to test the internal reliability of the newly developed scales and the procedure for the main study. Item-level analysis was used to detect problematic items and improve new scales.

The main finding relates to the “difficulty” of the newly developed scales; Saudi Arabian students agreed with most statements they were presented with. They may simply have very positive attitudes towards English and many Action Plans; developing questionnaires that are attuned better to the presence of latent factors within Saudi Arabian students is the main suggestion stemming from the pilot study. More specifically, this means coming up with more difficult items on both scales tested in the pilot study.

The exploratory factor analysis of Action Plans and Attitude towards English revealed that these are unidimensional constructs. It was expected that two factors would underlie the scores on Attitude towards English because the items were constructed considering two important components of attitudes (i.e., cognitive and affective). However, English attitude as operationalized in this study is a unitary construct. This goes against previous findings that a tripartite construct was more appropriate (Tseng, 2010; Lasagabaster, 2017) and can be at least partially explained by the differences between the questionnaires used by Tseng (2010) and Lasagabaster (2017) and the Attitude towards English questionnaire used in this study.

## 3 Main study

### 3.1 Method

#### 3.1.1 Research design

This is a correlational study, aiming to analyze the relationships between L2MSS, Action Plans, attitude towards English and English proficiency.

### 3.1.2 Participants

For the main study, 604 students (aged 17–33; 56.8% females) were recruited via convenience sampling. All were first-year students at their respective colleges in Saudi Arabia and enrolled in compulsory English courses. 34% studied Business Management at the time, while 66% studied Accounting. An English test based on TOEIC (Test of English for International Communication) was administered to participants (see “Instruments and procedure” below), the results showing that English proficiency among participants was fairly low. After discarding incomplete surveys or those with response patterns indicating lack of attention (e.g., the same response to all questions), as well as outliers, 488 participants’ data were included. The sample size can be considered to be in accordance with suggestions by Wolf et al. (2013) for Confirmatory Factor Analysis (CFA) and the requirements of IRT (Sahin and Anil, 2017).

### 3.1.3 Instruments and procedure

The L2MSS questionnaire was taken from Moskovsky et al. (2016), who sought to test Dörnyei’s theory in the Saudi context. They produced a version of the questionnaire adapted to the Saudi Arabian context with some new items (e.g., “Every Muslim should be able to speak English”). This questionnaire has 48 items across four subscales: ideal self (10 items), ought-to self (15 items), L2LE (15 items), and ILE (8 items)—a total of 48 items. The L2LE subscale has six inversely formulated items.

L2MSS allows us to calculate total score as a sum of all valid answers and calculate scores for subscales. Moskovsky et al.’s (2016) L2MSS was originally written in English. The researcher, a native Saudi Arabian, translated the questionnaire into Arabic, and an independent associate, an English professor in Saudi Arabia, provided a second translation. The two translations were considered together, forming the final “forward” translation. Two independent English literature students in Saudi Arabia then provided the “backward” translation where the English version was compared with the original English version of the questionnaire, allowing for further adjustments and improvements of the Saudi Arabian version.

A brief English test, consisting of 19 discrete tasks, was developed from the Reading part of the TOEIC practice test. These are typical incomplete sentence tasks; for instance, “Everyone in the department is expected to \_\_\_\_\_ the meeting,” with four possible answers: “Come,” “Attend,” “Be present,” or “Stay.” The final score is calculated as the sum of correct answers (19 = maximum). This is a fairly short test and furthermore it only captures one aspect of English proficiency. The decision to administer such a test was made due to time constraints of the study: namely, due to limited availability of participants, all research material had to be administered in one session.

The Action Plans and Attitude towards English instruments were described in the pilot study section. Final versions of these questionnaires, with 13 and 12 items respectively, were used in the main study. The final version of Attitude towards English has four inversely formulated items.

The procedure for the main study was the same as the pilot study procedure. Data was obtained during December 2021.

## 3.2 Analyses

Analysis of item performance (classical model and IRT) was performed for all instruments used in the study, as well as reliability

analyses of scales. Intercorrelations between L2MSS, Action Plans, English attitude, and English test performance were examined. Furthermore, the factor structure of L2MSS was also tested. CFA of Action Plans and English attitude was conducted in the AMOS (version 22) within SPSS.

## 3.3 Results

### 3.3.1 Item-level analysis

A classical item-level analysis of the L2MSS subscales revealed that ideal self and ought-to self-items were reliable and valid indicators of their respective latent factors. However, there were certain issues with some L2LE items (Table 2):

Notably, all items with low validity were inverse-formulated. However, the fact that not all inverse items were unreliable and invalid shows that this is not a consequence of our participants’ response sets or their lack of attention. Regarding homogeneity, the item “To be honest, I really have little interest in my English class” (Learnexp08) was the worst. This item seems unrelated to the overarching scale—L2LE—especially when compared with a more valid item such as “The English books that we use are really boring” (Learnexp14). This item obviously relates to L2LE, specifically the availability and relevance of learning materials. “I’m losing any desire I ever had to know English” (Learnexp10), another item with low validity, points to a general attitude towards English more clearly than indicating actual, hands-on learning experience. Thus, items Learnexp08 and Learnexp10 should be removed and replaced with more valid alternatives.

A different explanation is needed for the relatively low homogeneity (validity) of “I am sometimes worried that the other students in class will laugh at me when I speak English” (Learnexp12), and “It worries me that other students in my class seem to speak English better than I do” (Learnexp15). These items relate to anxiety and are likely closer to the construct of foreign language anxiety than to L2LE.

TABLE 2 Item-level metrics (classical model) for L2LE scale.

Item	REP	REL	H	B
Learnexp01	0.9593	4.986	7.413	6.372
Learnexp02	0.9565	0.4035	6.687	0.5879
Learnexp03	0.9583	0.4682	0.7011	0.6374
Learnexp04	0.9544	0.4446	0.4673	0.4547
Learnexp05	0.9646	0.4908	0.7014	0.5833
Learnexp06	0.9653	0.5536	0.7528	0.6280
Learnexp07	0.9466	0.5075	0.7121	0.6233
Learnexp08	0.8955	0.2935	0.0135	0.2112
Learnexp09	0.9541	0.4741	0.7151	0.6137
Learnexp10	0.9434	0.3962	0.2956	0.4840
Learnexp11	0.9599	0.3023	0.5686	0.4931
Learnexp12	0.8535	0.3829	0.1005	0.1988
Learnexp13	0.9170	0.4832	0.3319	0.4974
Learnexp14	0.9131	0.4457	0.3348	0.5084
Learnexp15	0.8707	0.3457	0.0596	0.2497

The classical analysis of intended learning behavior items did not reveal any issues and the same is true for Action Plans—a significant improvement compared to the initial scale version. Attitudes towards English items (Table 3) that were problematic were: “English pronunciation is so illogical that it drives me crazy” (Att12) and “The English language is too complex for me” (Att5). These items’ validity is questionable, with low homogeneity coefficients, meaning they are not ideal indicators of the attitude towards English; perhaps they indicate English language anxiety. The relevance of “English language is very boring” (Att9) and “English language is hard to learn” (Att3) is also questionable.

The IRT analysis Action Plans revealed that generally speaking items are “too easy”—even a relatively low presence of psychological constructs could trigger endorsement of a particular statement.

Winsteps provides “maps” of variables (items) and persons. These allow us to distinguish between “easy” and “difficult” items. This is crucial for future adjustments of the questionnaire, especially regarding specific problems in the Saudi context. “Whenever I think of my future career, I imagine myself using English,” and “The things I want to do in the future require me to use English,” were endorsed even by participants who had low Ideal self. These can be understood not as indicators of Ideal self but as general statements about the need to learn English; indeed, knowledge of English is necessary in most high-paying workplaces globally. We can contrast these items with “I can imagine myself studying in a university where all my courses are taught in English (maybe abroad in the future)” which is more “difficult.”

The easiest items of the ought-to self scale are: “Without learning English it will be difficult to travel to English speaking countries” (Oughtto2) and “I have to study English because I do not want to get bad marks in it” (Oughtto5). The relevance of Oughtto2 is questionable, while Oughtto5 is easy to agree with. The infit and outfit for all ought-to-self items fell between 0.5 and 1.5, but were higher for Oughtto2 (Infit = 1.41; Error SD = 5.1; Outfit = 1.6; Error SD = 6.2).

The L2LE items were somewhat better suited to the distribution of this variable in the population. However, higher levels of L2LE were not well covered by items, while the opposite was true for low levels. The worst item concerning infit and outfit is Learnexp12: “I

am sometimes worried that the other students in class will laugh at me when I speak English” (Infit = 1.41; Error SD = 6.8; Outfit = 1.49; Error SD = 7.9). It was also problematic regarding classical item-level metrics. Intended learning behavior items have solid infit and outfit, although more “difficult” items covering the variability of intended learning behavior in the upper realm were also absent.

Action Plans items fared well, compared to the initial pilot version. As expected, Item 13 (“I try to emulate the style of an English writer whose books I have read before”) was the most difficult item, still far away from the average presence of Action Plans among participants. Items 11 and 12, introduced after the pilot study, behaved as expected and were more difficult than most other items. However, these items seemed to cover the same portion of the distribution of Action Plans as Items 6 and 3; therefore, four items are suited to the same level of presence of Action Plans, while other levels are underrepresented.

Attitudes towards English items whose infit and outfit were under 0.7 or over 1.3 were 12, 8, 10, and 7. Thus, Item 12—“English pronunciation is so illogical that it drives me crazy”—was a bad addition. However, Item 11—“English language is more beautiful than other languages”—was a good addition. Item 4, which was changed into an item that only those with extremely positive attitudes towards English would agree with, was one of the easiest in the new version. Even participants who had negative attitudes toward English often agreed with “I want to know everything about the logic behind the English language.” This indicates how easy it was for Saudi participants to agree with the statements; it is very unlikely that most participants really want to know everything about English language logic.

### 3.3.2 Scale analyses

Table 4 shows descriptive statistics for all scales:

The normality analyses showed that only the total L2MSS scores taken as a whole do not deviate from normal distribution ( $KMS = 0.4$ ;  $p = 0.63$ ), while all the other subscales deviate significantly. Frequency histograms showed that ideal self, ought-to self, and ILEs were clearly negatively asymmetrical, reflecting the overall tendency of participants to agree with L2MSS items. However, L2LE scores had a slightly positively asymmetrical distribution, meaning that participants were less likely to agree with the items on this scale, compared to other scales. Action Plans and Attitude towards English also deviated from the normal distribution; they were both slightly leptokurtic—middle values were frequent, with Action Plans being skewed towards higher values. Moreover, the distribution of English language test scores was extremely positively asymmetrical, reflecting poor performance in this test. The reliability of the scales is provided in Table 5. Reliability estimates ranged from 0.789 (L2LE) and 0.932 (L2MSS), all at a satisfactory level.

Although the relationships among the L2MSS subscales, Action Plans, and Attitude towards English are significant and strong, there are no significant correlations among these variables and English language test scores. It seems that all the scales and subscales used are unrelated to English language performance: see correlations in Table 6.

The convergent validity of Attitude towards English and Action Plans is showcased in relatively high correlations with L2MSS (0.687 and 0.626, respectively). Exploring correlations between the two new scales and L2MSS factors allows us to further analyze their validity; specifically, the Action Plans scale is more weakly associated with L2LE (0.336) compared to Ought-to-self (0.582) and Intended

TABLE 3 Item-level metrics (classical model) for Attitude towards English scale.

Item	REP	REL	H	B
Att1	0.9603	0.2830	0.5780	0.5391
Att2	0.9712	0.4199	0.7110	0.6372
Att3	0.9475	0.5595	0.2306	0.4850
Att4	0.9683	0.3943	0.6792	0.5990
Att5	0.9329	0.6265	0.1547	0.4345
Att6	0.9528	0.5292	0.6765	0.5589
Att7	0.9745	0.5918	0.8231	0.7417
Att8	0.9730	0.5440	0.7923	0.7056
Att9	0.9443	0.5211	0.2864	0.5264
Att10	0.9692	0.5404	0.7831	0.7475
Att11	0.9516	0.5422	0.6964	0.5858
Att12	0.9272	0.3546	0.0633	0.3312



TABLE 4 Descriptive statistics for all scales.

	L2MSS total	Ideal self	Ought-to self	L2LE	ILEs	Action plans	English attitude	English test
Mean	178.3	39.9	56.1	49.9	32.3	50.7	41.5	5.2
SD	24.1	6.6	9.9	7.2	5.3	8.2	7.3	2.4
Range	98–236	15–50	24–75	28–71	11–40	12–60	12–60	1–18
Skewness (SE)	−0.441	−0.567	−0.215	0.211	−0.781	−0.488	0.49	1.268
Kurtosis	1.683	0.094	−0.099	0.173	0.526	1.257	0.814	3.648

TABLE 5 Reliability and representativeness of all scales.

	L2MSS total	Ideal self	Ought-to self	L2LE	ILEs	Action plans	English attitude
Cronbach's alpha	0.932	0.871	0.881	0.789	0.849	0.899	0.805
KMO representativeness	0.989	0.962	0.96	0.942	0.947	0.977	0.959

TABLE 6 Scales intercorrelations.

	L2MSS total	Ideal self	Ought-to self	L2LE	ILEs	Action plans	English attitude	English test
L2MSS		0.893**	0.859**	0.696**	0.886**	0.626**	0.687**	0.035
Ideal self			0.719**	0.515**	0.775**	0.571**	0.595**	0.021
Ought-to self				0.327**	0.703**	0.582**	0.486**	0.043
L2LE					0.551**	0.336*	0.617**	0.033
Intend. learn. eff.						0.590**	0.636**	0.009
Action Plans							0.510**	−0.005
Attitude								0.082
English test								

\* $p < 0.05$ ; \*\* $p < 0.01$ .

learning efforts (0.590), which is something that we would expect based on the nature of these factors. English attitude is more weakly associated with Ought-to-self (0.486) compared to Ideal self (0.595) and L2L2 (0.617). Exploratory factor analysis of L2MSS scores (ideal, ought-to selves, L2LE, and intended efforts) with maximum likelihood extraction and direct oblimin rotation, with a fixed number of four factors, yielded one main factor—a general L2MSS factor (accounting for 29% of extracted variance)—while the other three factors accounted for 7, 3, and 2.7% of total variance. All factors together accounted for 42% of extracted variance.

Factor 1 was labelled as a general L2MSS factor, after inspecting the structure and pattern matrices. The structure matrix showed that Factor 1 had relatively high loadings on most items of all L2MSS subscales. Pattern matrices were instrumental in interpreting the other three factors. Factor 2 mostly accounted for ought-to self and L2LE items, and this is also true for Factor 3. Factor 4 mainly had unique loadings on ideal-self items. Only Factor 1 had high pattern matrix loadings on intended learning behaviour items. See factor correlations in Table 7.

Factors 2, 3, and 4 are negatives of the original scales, indicating low ought-to self, negative L2 experience, and low intended efforts. Here, the chi-square for model fit provided by SPSS was significant (chi-square = 2,391,  $p = 0.000$ ), possibly meaning that more factors are needed for a better fit. The normality issues outlined above and

multivariate normality issues may have increased the significance of the chi-squared test of model fit.

Based on the factor analysis results of the pilot study, CFA was conducted with Attitude towards English and Action Plans scores. It was expected that Attitude towards English, as well as Action Plans, would converge towards a one-factor solution. The chi-square (1737,37;  $df = 274$ ) was statistically significant. The CFI was 0.735, with RMSE = 0.105. Considering recommendations available in AMOS, all three estimates indicated a relatively poor fit of the model.

Complete versions of all questionnaires, along with the English test used in the study, are provided as [Supplementary material](#).

### 3.4 Discussion

The study's hypotheses concerning the relatively poor criterion validity of L2MSS were proven. More specifically, L2MSS in our study is unrelated to performance on an English test. There are other studies which have also failed to identify an association between L2MSS components and English proficiency (Moskovsky et al., 2016; Al-Hoorie, 2018) while Alqahtani (2020) failed to detect a significant relationship between L2MSS and vocabulary size. It is possible that L2MSS does not relate to important and relevant real-life outcomes,

TABLE 7 Factor correlations.

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1		-0.123	-0.451	-0.586
Factor 2			0.107	0.002
Factor 3				0.321
Factor 4				

such as performance on an English test, although the failure to identify connections between L2MSS and English test performance in our study could be ascribed to the peculiarity of the Saudi context and the questionnaire issues stemming from this.

Action Plans and Attitude toward English did not predict English test performance, contrary to expectations. The factor structure of the L2MSS scale somewhat aligns with Dörnyei's theory. Namely, factor analysis identified a general L2MSS factor (Factor 1), alongside more specific Factors 2 and 3 (ought-to-self and L2LE), while Factor 4 accounted for ideal-self items. It has to be emphasized, however, that overall convergent validity of L2MSS has been proven in numerous studies, with L2MSS components being associated with other L2 acquisition constructs such as L2WTC (Amirian et al., 2021, p. 12; Li and Liu, 2021; Ebn-Abbasi et al., 2022; Zhou, 2022), growth mindset and academic engagement (Sadoughi et al., 2023), L2 academic resilience (perseverance) (Çelen, 2020), and L2 anxiety (Sadoughi and Hejazi, 2023). However, the connection between L2MSS and actual academic achievement remains elusive; findings of Sadoughi et al. (2023) relating to the association between ideal L2 self and academic engagement in this respect might be promising, although we have to keep in mind that academic engagement is not the same concept as actual academic achievement. Finally, it is possible that reconceptualization of L2MSS, for instance the introduction of 'feared L2 self' (Peker, 2020), could increase the criterion validity of L2MSS. Based on item-level analysis of L2MSS items, we can offer a set of suggestions regarding elimination or adjustment of certain L2MSS items: Learnexp08 and Learnexp10, as well as Learnexp12 and Learnexp15.

Although Action Plans and Attitude towards English converged towards one-factor solutions in the pilot study, the CFA of the same scales within the main study showed otherwise; namely, one-factor solutions of both Attitude towards English and Action Plans items were statistically invalid. This result is probably due to multivariate normality issues affecting the CFA.

One of the main issues, and one which complicates the interpretation of results, is the lack of both univariate and multivariate normality. Only L2MSS total scores were normally distributed, while the subscales of L2MSS, Action Plans, Attitude toward English, and English test scores all deviated from the normal distribution in various ways. It is likely that the peculiarity of the sample is one of the causes of these deviations; Saudi students are a very specific group for whom English language knowledge is highly desirable. With the exponential economic development of Saudi Arabia, gradual opening up to the world, and recent educational reforms and policy changes (Khawaji, 2022), English language proficiency has become an in-demand asset in Saudi Arabia, contributing to the participants' high scores on most scales in this study. The findings can therefore only be understood in the context of the rising importance of the English language in

Saudi Arabia. With Saudi Arabia's economic rise came increased global interconnectedness, necessitating knowledge of English, especially in the business sector. Thus, Saudi students must learn English to improve their own country. This is the main reason for our participants agreeing with most items. Interestingly, this is possibly another example of high conformism that characterizes Saudi Arabia (Gahwaji, 2008; Al-Rasheed, 2013; Al Lily and Alhazmi, 2017). For instance, scholars found that Saudi students have high conformity with the Communicative Language Teaching approach (Wajid and Saleem, 2017) despite having continually low English proficiency. High scores on L2MSS, Action Plans, and Attitude towards English, in the Saudi context, may be substantially affected by an urge to conform to society's demands, rather than due to a latent psychological construct.

We should also mention that participants' responses were likely affected by the study procedure. All questionnaires (L2MSS—48 items, Action Plans—13 items, attitude towards English—12 items) and English proficiency test (19 questions) were responded to in a single session due to time constraints. It would be better to present these questionnaires in 2 or 3 sessions. This probably affected participants' answers—for instance, a number of participants gave the same answer to all items (these responses were removed), possibly indicating study fatigue and frustration with the study protocol.

Another important finding relates to low English test scores. The test used in this study is based on TOEIC, which is a world-renowned English proficiency test used to assess all knowledge levels (from A1 to C1 in the Common European Framework of Reference). It was crucial to use a valid measure of English proficiency because there are concerns surrounding English course grading in Saudi Arabia, where numerous students pass their courses without acquiring the necessary knowledge (Barnawi and Al-Hawsawi, 2017).

Although the theoretical mean of the test is 9 correct answers, our participants' mean was 5.2, with very low variability ( $SD = 2.4$ ); the English test was simply too hard for these students. This was also the case in Moskovsky et al.'s (2016) study, in which an adaptation of IELTS was used to study L2MSS in the Saudi context. In this study, each question in the English test had four possible answers, meaning there was a 25% chance of a "lucky guess" for each incomplete sentence task (19 in total); indeed, one possible interpretation of the mean final score of 5.2 out of 19 is that most participants could have simply guessed the correct answers. Thus, the English test used in this study may not have accounted for the variability of knowledge among participants—reflected in turn in insignificant correlations between test scores and other variables. A logical conclusion would be that an easier test should be used in future studies of L2MSS in the Saudi context. Moreover, it would be useful to employ a more diverse English test in future studies, with a larger and more diverse set of tasks. Only one task type (sentence completion) was used in this study. All research materials had to be administered in one session due to time constraints (limited availability of participants), which led to making the English test as short as possible. Future studies should aim to assess English proficiency with at least a few types of different tasks (reading comprehension, essay writing, spelling, etc.) in order to account better for the diversity of English proficiency.

Keeping all this in mind, we should also note that motivation is not the only possible determinant of behavior and performance.

Intelligence and concrete abilities are also crucial factors for explaining language performance. Whether we consider language ability as part of general intelligence, or an independent ability, the relationships with real-life outcomes such as academic performance are strong and significant (Wesche et al., 1982; Sasaki, 1993; Saricaoglu and Arikan, 2009). This may be a more promising approach, at least if we are concerned with more direct predictors of language performance. Research into motivational factors, however, is important even if it will not help improve language performance. By pointing to specific components of motivation and their relationships, motivation research can help enhance students' overall learning experience and satisfaction with foreign language courses, making them, for instance, more likely to use a foreign language in a wider variety of circumstances. Thus, L2MSS could possibly benefit from the introduction of language ability factors which, by definition, are related to real-life outcomes such as performance on an English test. It would be interesting to explore the relationships between L2MSS and language ability factors, especially in younger schoolchildren, and whether language ability can be fostered early on by improving L2 motivation.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

## References

- Ajzen, I. (2005). *Attitudes, personality and behaviour*. McGraw-Hill education: UK.
- Al Lily, A. E., and Alhazmi, A. A. (2017). Passive conformism in academia: Saudi organization, education, and technology. *Digest Middle East Stud.* 26, 340–361. doi: 10.1111/dome.12109
- Al-Hoorie, A. H. (2018). The L2 motivational self system: a meta-analysis. *Stud. Second Lang. Learn. Teach.* 8, 721–754. doi: 10.14746/ssl.2018.8.4.2
- Almuhaimeed, S. A. (2002). Evaluating the validity of Saudi English language undergraduate students' results in light of proposed criteria. *J. lang. linguist. stud.* 18(special issue 1), 359–378. Available at: <https://www.jlls.org/index.php/jlls/article/view/3818/1013>.
- Alqahtani, A. F. (2020). The relationship between the Saudi Cadets' learning motivation and their vocabulary knowledge. *Engl. Lang. Teach.* 13, 1–10. doi: 10.5539/elt.v13n4p1
- Arabai, F. (2016). Factors underlying low achievement of Saudi EFL learners. *Int. J. English Linguist.* 6, 21–37. doi: 10.5539/ijel.v6n3p21
- Al-Rasheed, M. (2013). *A most masculine state: Gender, politics and religion in Saudi Arabia* (No. 43). Cambridge: Cambridge University Press.
- Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. *Int. J. English Linguist.* 4, 17–26. doi: 10.5539/ijel.v4n5p17
- Alshammari, H. A. (2022). Investigating the low English proficiency of Saudi EFL learners. *Arab World English J.* 13, 129–144. doi: 10.24093/awej/vol13no1.9
- Amirian, Z., Shayanfar, M., and Rezazadeh, M. (2021). Structural equation model of Teachers' mindsets, self-efficacy, and emotional experiences and Iranian EFL Learners' willingness to communicate and L2 motivational self system. *Two Q. J. English Lang. Teach. Learn. Univ. Tabriz* 13, 1–17. doi: 10.22034/ELT.2021.48218.2448
- Bagozzi, R. P., Tybout, A. M., Craig, C. S., and Sternthal, B. (1979). The construct validity of the tripartite classification of attitudes. *J. Mark. Res.* 16, 88–95. doi: 10.1177/002224377901600113
- Barnawi, O. Z., and Al-Hawsawi, S. (2017). "English education policy in Saudi Arabia: English language education policy in the Kingdom of Saudi Arabia: current trends, issues and challenges" in *English language education policy in the Middle East and North Africa*. ed. R. Kirkpatrick (Cham: Springer), 199–222.
- Boo, Z., Dörnyei, Z., and Ryan, S. (2015). L2 motivation research 2005–2014: understanding a publication surge and a changing landscape. *System* 55, 145–157. doi: 10.1016/j.system.2015.10.006
- Çelen, E. (2020). A structural equation model on EFL tertiary level Students' academic buoyancy, academic resilience, Reconceptualized L2 motivational self system, and their academic achievement. Doctoral dissertation, Bilkent Üniversitesi (Turkey).
- Csizér, K. (2019). "The L2 motivational self system" in *The Palgrave handbook of motivation for language learning*. eds. M. Lamb, K. Csizér, H. Alastair and R. Stephen. (Cham: Springer). 71–93.
- Dörnyei, Z. (2009). "The L2 motivational self system" in *Motivation, language identity and the L2 self*. eds. Z. Dörnyei and E. Ushioda. (Bristol: Multilingual Matters). 9–42.
- Dörnyei, Z. (2019). Towards a better understanding of the L2 learning experience, the Cinderella of the L2 motivational self system. *Stud. Second Lang. Learn. Teach.* 9, 19–30. doi: 10.14746/ssl.2019.9.1.2
- Ebn-Abbasi, F., Nushi, M., and Fattahi, N. (2022). The role of L2 motivational self system and grit in EFL learners' willingness to communicate: a study of public school vs. private English language institute learners. *Front. Educ.* 7:837714. doi: 10.3389/feduc.2022.837714
- Eusafzai, H. A. K. (2013). L2 motivational selves of Saudi preparatory year EFL learners: a quantitative study. *Arab World English J.* 4, 183–201. Available at: <https://awej.org/images/AllIssues/Volume4/Volume4Number4Dec2013/11.pdf>.
- Fajgelj, S., and Kosanović, B. (2001). Nova i stara ajtem analiza-poređenje (New and Old Item Analysis). *Psihologija* 1, 83–110. Available at: <https://scindeks-clanci.ceon.rs/data/pdf/0048-5705/2001/0048-57050101083F.pdf>.
- Faruk, S. M. G. (2013). English language teaching in Saudi Arabia—a world system perspective. *Buletinul Stiintific al Universitatii Politehnica din Timisoara, Seria Limbi Moderne*. Available at: [https://www.academia.edu/download/35823341/06\\_English\\_Language\\_Teaching\\_in\\_Saudi\\_Arabia.pdf](https://www.academia.edu/download/35823341/06_English_Language_Teaching_in_Saudi_Arabia.pdf).
- Fatimah, S. (2015). A study of students' performance in calculus and their attitudes toward the course using tripartite model. *Int. J. Humanit. Soc. Sci.* 1, 30–35. doi: 10.20469/ijhss.20005
- Gahwaji, N. M. J. (2008). Learning through play for preschool children in Jeddah—opportunity and conformity. (project), King Abdulaziz university. Available at: [https://www.kau.edu.sa/Show\\_Res.aspx?Site\\_ID=0003304&Lng=EN&RN=25147](https://www.kau.edu.sa/Show_Res.aspx?Site_ID=0003304&Lng=EN&RN=25147).
- Gardner, R. C. (2001). "Integrative motivation and second language acquisition" in *Motivation and second language acquisition*. eds. Z. Dörnyei and R. Schmidt (Honolulu, HI: University of Hawaii Press), 1–19.
- Hamouda, A. (2013). An exploration of causes of Saudi students' reluctance to participate in the English language classroom. *Int. J. English Language Education.* 1, 17–34. doi: 10.5296/ijele.v1i1.2652

## Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

## Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2024.1247701/full#supplementary-material>

- He, T. H., Chang, S. M., and Chen, S. H. E. (2011). Multiple goals, writing strategies, and written outcomes for college students learning English as a second language. *Percept. Mot. Skills* 112, 401–416. doi: 10.2466/11.21.28.PMS.112.2.401-416
- Kaiser, F. G., and Wilson, M. (2019). The Campbell paradigm as a behavior-predictive reinterpretation of the classical tripartite model of attitudes. *Eur. Psychol.* 24, 359–374. doi: 10.1027/1016-9040/a000364
- Kang, D. (2017). An exploratory factor analysis of L2 motivational self system in Korean EFL college contexts. *Foreign Lang. Educ.* 24, 1–21. doi: 10.15334/FLE.2017.24.1.1
- Khawaji, A. (2022). Transition of English language teaching in Saudi Arabia: a critical evaluative study. *Arab world. Engl. J.* 13, 265–280. doi: 10.24093/awej/vol13no4.17
- Kim, T. Y., and Kim, Y. (2021). Structural relationship between L2 learning motivation and resilience and their impact on motivated behaviour and L2 proficiency. *J. Psycholinguist. Res.* 50, 417–436. doi: 10.1007/s10936-020-09721-8
- Kim, K. H., Lu, J., and Estrada-Hernandez, N. (2015). Attitudes toward people with disabilities: the tripartite model, social desirability, and other controversial variables. *J. Asia Pacific Counsel.* 5, 23–37. doi: 10.18401/2015.5.1.2.
- Lake, J. (2013). Positive L2 self: linking positive psychology with L2 motivation. In M. T. Apple, SilvaD. Da and T. Fellner (Eds.), *Language learning motivation in Japan* (pp. 225–244). Bristol, UK: Multilingual Matters.
- Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Lang. Learn.* 62, 997–1023. doi: 10.1111/j.1467-9922.2012.00719.x
- Lasagabaster, D. (2017). Language learning motivation and language attitudes in multilingual Spain from an international perspective. *Mod. Lang. J.* 101, 583–596. doi: 10.1111/modl.12414
- Lee, J. S., and Lee, K. (2021). The role of informal digital learning of English and L2 motivational self system in foreign language enjoyment. *Br. J. Educ. Technol.* 52, 358–373. doi: 10.1111/bjet.12955
- Li, T., and Liu, Z. (2021). Exploring effects of the second language motivational self system on Chinese EFL learners' willingness to communicate in English and implications for L2 education. *J. Higher Educ. Res.* 2, 169–177. doi: 10.32629/jher.v2i4.404
- Linacre, J. M., and Wright, B. D. (2000) Winsteps from <http://www.Winsteps.Com/index.htm>.
- Moskovsky, C., Assulaimani, T., Racheva, S., and Harkins, J. (2016). The L2 motivational self system and L2 achievement: a study of Saudi EFL learners. *Mod. Lang. J.* 100, 641–654. doi: 10.1111/modl.12340
- Ostrom, T. M. (1969). The relationship between the affective, behavioural, and cognitive components of attitude. *J. Exp. Soc. Psychol.* 5, 12–30. doi: 10.1016/0022-1031(69)90003-1
- Peker, H. (2020). Feared L2 self as an emerging component of the reconceptualized L2 motivational self system. *Eur. J. Appl. Linguist.* 6, 361–386. doi: 10.32601/ejal.834658
- Rasch, G. (1960). *Studies in mathematical psychology: Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.
- Sadoughi, M., and Hejazi, S. Y. (2023). How can L2 motivational self system enhance willingness to communicate? The contribution of foreign language enjoyment and anxiety. *Curr. Psychol.* 42, 1–13. doi: 10.1007/s12144-023-04479-3
- Sadoughi, M., Hejazi, S. Y., and Lou, N. M. (2023). How do growth mindsets contribute to academic engagement in L2 classes? The mediating and moderating roles of the L2 motivational self system. *Soc. Psychol. Educ.* 26, 241–261. doi: 10.1007/s11218-022-09744-0
- Sahin, A., and Anil, D. (2017). The effects of test length and sample size on item parameters in item response theory. *Educ. Pract. Theory* 17, 321–335. doi: 10.12738/estp.2017.1.0270
- Saricaoglu, A., and Arikan, A. (2009). A study of multiple intelligences, foreign language success and some selected variables. Available at: <https://eric.ed.gov/?id=ED506218>.
- Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: a structural equation modeling approach. *Lang. Learn.* 43, 313–344. doi: 10.1111/j.1467-1770.1993.tb00617.x
- Smith, A. N. (1971). The importance of attitude in foreign language learning. *Modern Lang. J.* 55, 82–88. doi: 10.2307/321854
- Suhr, D. D. (2005). Principal component analysis vs. exploratory factor analysis. Available at: <https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/203-30.pdf>.
- Thompson, A. S., and Erdil-Moody, Z. (2016). Operationalizing multilingualism: language learning motivation in Turkey. *Int. J. Biling. Educ. Biling.* 19, 314–331. doi: 10.1080/13670050.2014.985631
- Tseng, W. T. (2010). Multidimensionality and hierarchical structure of the English vocabulary learning attitude scale. *Soc. Behav. Pers.* 38, 907–918. doi: 10.2224/sbp.2010.38.7.907
- Wajid, M. A., and Saleem, M. (2017). Learner conformity to communicative language teaching approach in EFL contexts: a case study in Saudi Arabia. *Int. J. Lang. Linguist.* 4, 240–249. Available at: [https://ijllnet.com/journals/Vol\\_4\\_No\\_4\\_December\\_2017/26.pdf](https://ijllnet.com/journals/Vol_4_No_4_December_2017/26.pdf).
- Wesche, M., Edwards, H., and Wells, W. (1982). Foreign language aptitude and intelligence. *Appl. Psycholinguist.* 3, 127–140. doi: 10.1017/S0142716400006664
- Wolf, E. J., Harrington, K. M., Clark, S. L., and Miller, M. W. (2013). Sample size requirements for structural equation models: an evaluation of power, Bias, and solution propriety. *Educ. Psychol. Meas.* 76, 913–934. doi: 10.1177/0013164413495237
- Zhou, C. (2022). The relationship between Chinese EFL learners' L2 motivational self system and willingness to communicate. *Theory Pract. Lang. Stud.* 12, 2374–2380. doi: 10.17507/tpls.1211.18