

This is the peer-reviewed version of the following article: Loorbach, N. , Peters, O. , Karreman, J. and Steehouder, M. (2015), Validation of the Instructional Materials Motivation Survey (IMMS) in a self-directed instructional setting aimed at working with technology. *Br J Educ Technol*, 46: 204-218., ***which has been published in final form at [10.1111/bjet.12138](https://doi.org/10.1111/bjet.12138). This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.***

Validation of the Instructional Materials Motivation Survey (IMMS) in a self-directed instructional setting aimed at working with technology

Nicole Loorbach, Oscar Peters, Joyce Karreman and Michaël Steehouder

University of Twente, Faculty of Behavioural Sciences

Nicole Loorbach completed her PhD at the Department of Communication Science — Corporate and Marketing Communication at the University of Twente (The Netherlands). She is cofounder, research director and researcher at PURE Research, and founder of DUH TXT, where she creates and transforms texts to communicate what matters most. She also teaches others to do the same, in scientific, corporate and personal communication. Oscar Peters (PhD) is Director of Education at the School of Management and Governance at the University of Twente. Joyce Karreman is an assistant professor at the Department of Communication Science — Corporate and Marketing Communication at the University of Twente. She teaches courses in document design and user support. Michaël Steehouder is professor emeritus of technical communication at the University of Twente. He was chair of the Communication Department. Email: n.r.loorbach@utwente.nl

Abstract

The ARCS Model of Motivational Design has been used myriad times to design motivational instructions that focus on Attention, Relevance, Confidence, and Satisfaction in order to motivate students. The Instructional Materials Motivation Survey (IMMS) is a 36-item situational measure of people's reactions to instructional materials in the light of the ARCS Model. Although the IMMS has been used often, both as a pre-test and a post-test tool serving as either a motivational needs assessment prior to instruction or a measure of people's reactions to instructional materials afterwards, the IMMS so far has not been validated extensively, taking statistical and theoretical aspects of the survey into account. This article describes such an extensive validation study, for which the IMMS was used in a self-directed instructional setting aimed at working with technology (a cellular telephone). Results of structural equation modeling (SEM) show that the IMMS can be

reduced to 12 items. This Reduced IMMS (RIMMS) is preferred over the original IMMS. The parsimonious RIMMS measures the four constructs Attention, Relevance, Confidence, and Satisfaction of the ARCS Model well, and reflects its conditional nature.

Introduction

In the field of educational science, the ARCS Model of Motivational Design (Keller, 1983, 1987, b, c, 1999, 2010; Keller & Kopp, 1987) has been used myriad times to apply motivational strategies to instructional materials, and to test their effects. Although the model was originally designed to influence student motivation in a classic learning setting, with face-to-face interaction between teacher and students, by now it has also been thoroughly applied to and tested in other settings like computer-assisted instruction, and computer-based and distance education (eg. Astleitner & Hufnagl, 2003; Bellon & Oates, 2002; Chang & Lehman, 2002; Chyung, Winiecki, & Fenner, 1999; Keller, 1999; Shellnut, Knowlton, & Savage, 1999; Song & Keller, 2001). And in recent years, the ARCS Model has been applied to and tested in self-directed, print-based instructional settings, applying it to cell phone user instructions and testing for effects on users likely to benefit from motivational instructions (see Lorbach, Karreman and Steehouder, 2007 and 2013, b for elaborate descriptions).

Keller's publications on the ARCS Model show a similar expansion of its scope as other publications over time. In his early work, Keller (1987, b, c) speaks of "students' motivation to learn," "education," "course," "lesson," and "classroom setting." In 1999, he states that "it is one thing to design for learner motivation in a classroom setting where teachers or facilitators can respond to changes as soon as they sense them. It is a greater challenge to make self-directed learning environments responsive to the motivational requirements of learners" (p. 39).

The ARCS Model of Motivational Design

The ARCS Model of Motivational Design is based on an extensive review of the motivational literature which led to a clustering of motivational concepts into four constructs: (A)ttention, (R)elevance, (C)onfidence, and (S)atisfaction (Keller, 2010, p. 44). According to Keller (2010, pp. 44—45), the following goals have to be met for people to be motivated to learn:

(A) People's curiosities and interests should be stimulated and sustained.

- (R) Before people can be motivated to learn, they will have to believe that the instruction is related to important personal goals or motives and feel connected to the setting.
- (C) Even if people believe the content is relevant and they are curious to learn it, they still might not be appropriately motivated due to too little or too much confidence, or expectancy for success. They could have well-established fears of the topic, skill, or situation that prevent them from learning effectively. Or, at the other extreme, they might believe incorrectly that they already know it and overlook important details in the learning activities.

Keller (2010, p. 46) states that being successful in achieving these first three motivational goals (attention, relevance, and confidence) results in people being motivated to learn.

- (S) To have a continuing desire to learn, people must have feelings of satisfaction with the process or results of the learning experience.

The above description of the ARCS Model is visualized in Figure 1.

[INSERT FIGURE 1 HERE]

Practical strategies described in the ARCS theory and in the Motivational Tactics Checklist (see Keller, 2010, pp. 287—291) can be used to achieve each of the four goals. To measure whether these goals have been met, and to measure learners' motivational needs prior to applying ARCS strategies, Keller (2010, pp. 277—286) designed the Instructional Materials Motivation Survey (IMMS), a 36-item situational measure of people's reactions to instructional materials in the light of the ARCS Model. As such, it measures people's scores on an attention, relevance, confidence, and satisfaction construct, cumulatively resulting in an overall motivation score.

Effects of motivational instructions in a self-directed instructional setting

A previous study on the effects of ARCS-based motivational instructions in a self-directed instructional setting tested for effects of three motivational manipulations in cell phone user instructions, respectively focusing on attention, relevance, and confidence (see Loorbach et al., 2007 for an elaborate description). Seventy-nine Dutch senior participants between 60 and 70 years of age filled out questionnaires and performed three tasks with a

cell phone, using either a control version or one of three motivational versions of the user instructions. Participants in this study were seniors, because they belong to a user group that is known for being less experienced with relatively new technology devices like cellular telephones (Schwender & Köhler, 2006), and are therefore more likely to benefit from motivational instructions.

Results showed that participants using either a version of the instructions that focused on relevance or a version that focused on confidence, performed more tasks correctly than participants in the control condition, using instructions without motivational manipulations. This study also showed positive effects of motivational instructions on behavior-deduced motivation. For this measure, we only included participants who did not complete the task, and we checked whether they felt too frustrated and gave up prematurely, or persisted and their efforts were stopped by the researcher after they had been working on the task for 15 minutes. Results showed that participants using the confidence-focused user instructions persisted in working on the third task, where they had to edit a contact's phone number, more often ($p < .05$). A tendency toward a similar effect existed for the first task, where they had to change the cell phone's ring tone ($p < .10$).

So even though the ARCS Model of Motivational Design was not originally designed to increase user motivation in self-directed instructional settings, its potential was discovered for such settings. Its potential was especially discovered concerning confidence-focused instructions, which positively affected participants' task performance and their persistence in trying to complete tasks. This is in line with the expectations of the ARCS Model: When it was first developed, Keller (1987c) stated that "differences in confidence, the third major component of the model, can influence a student's persistence and accomplishment" (p. 5). However, even though the behavior of participants using the control version and participants using the confidence version of the instructions statistically differed in persistence, these findings were nonexistent according to their motivation scores on the IMMS.

A possible explanation is that participants who used the motivational instructions did have an increased motivation level, but were not aware of it, and therefore a self-report measure like the IMMS did not pick up on it, even though their behavior showed otherwise. According to Song and Keller (2001), "the use of self-report methods for measuring motivation [is] limited in that such methods [require] students to indicate their perceived motivation level, which might have been different from their actual amount of effort – a

more accurate measure of motivational behavior” (p. 20). The latter is a confirmation of Keller’s words (1983): “*Effort* refers to whether the individual is engaged in actions aimed at accomplishing the task. Thus, effort is a direct indicator of motivation” (p. 391).

Another explanation for the discrepancy between outcomes of behavior-deduced and self-reported measures of motivation is that the IMMS might not be suitable for measuring motivational differences in our self-directed instructional setting after all. The ARCS Model—which the IMMS is based on—did prove effective in our setting (see Loorbach, Karreman, & Steehouder, 2007, 2013, b), in spite of its original aim to increase motivation in instructor-facilitated instructional settings (Keller, 2010, p. 310). This article describes a validation of the IMMS to rule out or confirm this second explanation; that the IMMS as it is might not be suitable for measuring motivational differences in a self-directed instructional setting.

Aim of study

Although the IMMS has been used often to measure people’s perceived motivation concerning motivationally adapted instructions (eg. Choi & Johnson, 2005; Means, Jonassen, & Dwyer, 1997), the survey has never been validated from A to Z, unraveling its statistical and theoretical strengths and weaknesses in relation to all aspects of the underlying ARCS theory. Keller (2010) validated his survey by offering undergraduate students one of two sets of instructional materials (a control set and a set focusing on all four goals of motivational instructions, namely *attention*, *relevance*, *confidence*, and *satisfaction*), and having them fill out the IMMS. Results showed that scores on the experimental instructions were significantly higher than for the control instructions. Keller (2010) refers to Naime-Diffenbach’s study (1991), which showed similar results: When there is actual variation in materials concerning the motivational dimensions of the ARCS Model, then the IMMS scores will reflect these differences. Keller also refers to Small and Gluck’s study (1994), which confirmed the four-component taxonomy of the ARCS theory. Even though the IMMS did not reflect the behavioral differences or the variation in materials in Loorbach et al.’s study (2007), the aforementioned studies do tell us that the IMMS has potential to function as a manipulation check when there is actual variation in materials. Still, these studies have not examined the validity of the IMMS survey with respect to both statistical and theoretical strengths and weaknesses in the light of the underlying ARCS theory.

Huang, Huang, Diefes-Dux and Imbrie (2006) have made a first attempt to validate the IMMS in such a way, in a computer-based tutorial setting, using structural equation modelling (SEM). Their study results suggest that 16 of the original items should be excluded from the IMMS. The remaining 20-item scale is supported statistically, but lacks theoretical support, since some of the remaining items no longer belong to the same constructs they were originally assigned to. Also, the reflection of the conditional nature of the ARCS theory in the IMMS was not considered.

Keller (2010), thereby referring to Huang et al. (2006), forewarns of applying traditional factor analysis to the IMMS and of obtaining its intended factor structure, since the four subscales can have high intercorrelations (p. 286). He explains that this is in part because the IMMS was designed to measure situation-specific attitudes and not psychological constructs. However, we feel that testing for model fit is called for when motivationally distinctive materials are used. Therefore, the study described in this article attempts to validate the IMMS using SEM, in two ways the IMMS was not originally developed for:

- (1) In a self-directed and therefore non-interactive setting, which focused on learning and performing instead of learning per se, and
- (2) With senior users instead of students.

The IMMS will be validated as a post-test tool, measuring people's reactions to instructional materials (as opposed to measuring their motivational needs prior to applying ARCS strategies). In a world where technological development skyrockets, and senior users especially are having trouble keeping up the pace (Chisnell & Redish, 2005), helping this user group to fully benefit their new products by providing motivational user instructions is essential. Having a tool that appropriately measures their motivational needs prior to designing motivational instructions, and that subsequently checks whether strategies to improve the instructional materials motivationally have had the desired effects on senior users is equally as essential.

In order to validate the IMMS in a self-directed instructional setting with people likely to benefit from motivational instructions, a longitudinal trend study was conducted among Dutch seniors between 60 and 70 years of age, who used cell phone user instructions with or without motivational manipulations. This article describes two studies. Study 1's data will be used to test for model fit of the original model: Do the IMMS scores confirm the four-component taxonomy of the ARCS theory? And do they reflect the conditional nature of the

ARCS Model (see Figure 2)? Subsequently, two alternative models will be tested, representing additional causal connections other than the ones suggested by the ARCS theory which could be plausible in the light of the same theory. The alternative model AC contains an additional causal connection between Attention and Confidence, and the alternative model RS contains an additional causal connection between Relevance and Satisfaction, when compared to the original model. Finally, the items of the best fitting model, represented in the Reduced Instructional Materials Motivation Survey (RIMMS), will be retested in Study 2.

[INSERT FIGURE 2 HERE]

Method

Participants and procedure

Seventy-nine Dutch seniors (39 males and 40 females; age range 60—70, $M = 65.68$, $SD = 3.09$) participated in Study 1 (see Loorbach et al., 2007), which tested for effects of motivational cell phone user instructions, on aspects of usability, motivation, and confidence. Study 2 was based on the effects of Study 1, and was therefore comparable to Study 1. This time, 59 seniors participated (30 males and 29 females; age range 60—70, $M = 65.54$, $SD = 2.81$). All 138 participants either replied to an advertisement in newsletters of several elderly associations, to a flyer that was put in their mail box, or to a request of an acquaintance who had already participated. Participants received monetary compensation for their cooperation. If they chose to, this compensation was transferred to the bank account of the elderly association the participant belonged to, or to a charitable institution of their choosing. The only selection criteria we used were that the participants were between 60 and 70 years of age, and that they did not have any prior knowledge on how to use the brand of cell phone used in the studies.

In both studies, participants filled out pre-test questionnaires, subsequently performed three tasks with the cell phone and one of four (Study 1) or one of three (Study 2) sets of user instructions, and then filled out post-test questionnaires, one of which was a translated and modified version of either the IMMS (Study 1, 36 items) or the RIMMS (Study 2, 12 items).

Materials

Instructional Materials Motivation Survey (IMMS)

Keller (2010) designed the IMMS to be in correspondence with the theoretical foundation represented by the motivational concepts and theories comprising the ARCS Model. The survey was designed to measure reactions to self-directed instructional materials. It is a situation-specific self-report measure that can be used to estimate learners' motivational attitudes in the context of virtually any delivery system. The IMMS can be used with print-based self-directed learning, computer-based instruction, or online courses that are primarily self-directed (p. 277).

The IMMS is a 36-item scale consisting of four subscales (cf. the ARCS Model's four constructs Attention, Relevance, Confidence, and Satisfaction). According to Keller (2010) "each of the four subscales can be used and scored independently" (p. 282). Scoring can also be done for the total scale measuring motivation. Since there are no norms for this survey, there are not any set numbers to indicate low or high scores.

In Study 1, a translated, adapted version of the IMMS was used in a print-based, self-directed instructional setting. Translation into Dutch was achieved using back-translation, and the IMMS was adapted to our setting, for example by replacing "this lesson" by "these user instructions." See Appendix A for the adapted version of the IMMS.

Reduced Instructional Materials Motivation Survey (RIMMS)

In Study 2, the RIMMS was used. This is a reduced version of the IMMS version that was used in Study 1. The RIMMS is a 12-item scale consisting of four subscales, each comprising 3 items. Scoring can be done for each subscale independently, or for the total scale measuring motivation. Asterisks in Appendix A indicate RIMMS items.

Data analysis

SEM analyses were conducted using Amos 19.0 with maximum likelihood estimation to test the hypothesized model. We applied Anderson and Gerbing's (1988) two-step approach by first testing the fit of our model's measurement components. Subsequently, when the fit of the measurement model was acceptable, the fit of the structural model was tested.

As suggested by Holbert and Stephenson (2002), the following model fit indices were used: The chi-square estimates with degrees of freedom as it is the most commonly used indice to make comparisons across models (Hoyle & Panter, 1995). The ratio between chi-square and

degrees of freedom should not exceed three for models with a good fit (Ullman, 2001). Additionally, the standardized root mean squared residual (SRMR) as a second absolute fit statistic (Hu & Bentler, 1999), in combination with the Tucker-Lewis index (TLI) as incremental index, and the root mean squared error of approximation (RMSEA; Browne & Cudeck, 1993) are reported. Hu and Bentler (1999) recommend using a cutoff value close to .95 for TLI in combination with a cutoff value close to .09 for SRMR to evaluate model fit, and the RMSEA close to .06 or less. The goodness-of-fit measures used in this study, to test the fit of the conceptual model with the empirical data, are among the measures least affected by sample size (Fan, Thompson, & Wang, 1999).

Results

Study 1: IMMS

Using a first-order confirmatory factor analysis, the measurement model estimated the extent to which the observed 36 IMMS items loaded onto their respective latent variables. All latent construct variables were allowed to covary with and predict all variables in the model. However, in line with Gerbing and Anderson's recommendations (1984), errors were not allowed to correlate when testing the models.

Measurement model

The initial measurement model with 36 items did not fit the data well: $\chi^2(588) = 1066.81$; $\chi^2/df = 1.81$; SRMR = .10; TLI = .61; RMSEA = .09 (90% confidence interval [CI] = .08, .10). Items with highly correlated error variance and items that loaded poorly onto its unique factor were removed. This procedure resulted in a reduction of the number of observed indicators of the latent constructs. The resulting modified measurement model with twelve IMMS items (forming the RIMMS) generated an adequate fit: $\chi^2(48) = 86.32$; $\chi^2/df = 1.80$; SRMR = .06; TLI = .90; RMSEA = .09 (CI = .06, .12). The internal consistency of the ARCS measures was above aspiration level ($\alpha > .70$), except for the Relevance construct ($\alpha > .68$). Although the internal consistency of Relevance was below aspiration level, it was at an acceptable level to be included in further analyses. The correlation matrix of the observed variables is shown in Table 1.

[INSERT TABLE 1 HERE]

Structural model

The results obtained from testing the validity of a causal structure of the hypothesized ARCS model showed an adequate fit: $\chi^2 (51) = 89.59$; $\chi^2 /df = 1.76$; SRMR = .09; TLI = .91; RMSEA = .09 (CI = .06, .12). Table 2 summarizes the mean and standard deviation, Cronbach's α , the factor loading (β), and the squared multiple correlation (R^2) of the observed indicators.

[INSERT TABLE 2 HERE]

Path model

The path model with standardized path coefficients is featured in Figure 3: The standardized path coefficients show a significant direct effect of Attention on Relevance, a significant direct effect of Relevance on Confidence, and a significant direct effect of Confidence on Satisfaction.

[INSERT FIGURE 3 HERE]

Alternative models

To test the two alternative models, two additional second-order confirmatory factor analyses were performed. The results obtained from testing the validity of a causal structure of both the alternative AC model ($\chi^2 [51] = 89.59$; $\chi^2 /df = 1.76$; SRMR = .09; TLI = .91; RMSEA = .09 [CI = .06, .12]) and the alternative RS model ($\chi^2 [51] = 89.59$; $\chi^2 /df = 1.76$; SRMR = .09; TLI = .91; RMSEA = .09 [CI = .06, .12]) showed a similar adequate fit as the ARCS model. The standardized path coefficients of the AC model are significant, except for the path from Attention to Confidence. Also, the standardized path coefficients of the RS model are significant, except for the path from Relevance to Satisfaction. The results obtained from testing the validity of the two alternative models indicate that the alternative models are equal to the original ARCS model.

Study 2: RIMMS

To test the validity of a causal structure of the hypothesized ARCS model, the sum score of each factor was obtained (see Table 3). All latent construct variables were allowed to covary with and predict all variables in the model, and errors were not allowed to correlate when testing the models. The internal consistency of the ARCS measures was well above the

aspiration level of .70 ($\alpha > .81$ for all constructs). The correlation matrix of the factors is shown in Table 3.

[INSERT TABLE 3 HERE]

Structural model

The results obtained from testing the validity of a causal structure of the hypothesized ARCS model showed a moderate fit, $\chi^2(3) = 9.73$; $\chi^2/df = 3.24$, SRMR = .04; TLI = .95; RMSEA = .20 (CI = .07, .34). Table 2 summarizes the mean and standard deviation, and Cronbach's α of the four factors.

Path model

The path model with standardized path coefficients is featured in Figure 4. The standardized path coefficients in Figure 4 show a significant direct effect of Attention on Relevance, a significant direct effect of Relevance on Confidence, and a significant direct effect of Confidence on Satisfaction.

[INSERT FIGURE 4 HERE]

Alternative models

To test the two alternative models, two additional second-order confirmatory factor analyses were performed. The results obtained from testing the validity of a causal structure of both the alternative AC model ($\chi^2(2) = 6.26$; $\chi^2/df = 3.13$; SRMR = .03; TLI = .95; RMSEA = .19 [CI = .03, .37]) and the alternative RS model ($\chi^2(2) = 6.87$; $\chi^2/df = 3.44$; SRMR = .03; TLI = .95; RMSEA = .21 [CI = .05, .38]) showed a similar moderate fit as the ARCS model. The standardized path coefficients of the AC model are significant, except for the path from Attention to Confidence. Also, the standardized path coefficients of the RS model are significant, except for the path from Relevance to Satisfaction. The results obtained from testing the validity of the two alternative models indicate that the alternative models are equal to the original ARCS model.

Conclusion and discussion

This validation study has shown that in a self-directed instructional setting aimed at motivating seniors to learn and perform tasks, the data of the 12-item measure RIMMS fit

the ARCS Model and its four constructs Attention, Relevance, Confidence, and Satisfaction better than the original 36-item measure IMMS. Also, the alternative models—with an additional causal relationship between either Attention and Confidence, or Relevance and Satisfaction—are statistically equally fitting to the original model, which adheres causal relationships according to the ARCS Model (see Figure 1). Since the alternative models do not improve model fit, and the original ARCS model is more parsimonious, the ARCS model is theoretically better fitting than the alternative models. In Keller’s words (2010), but replacing IMMS with RIMMS, this study has “confirmed the empirical validity of the [RIMMS] by confirming the four-component taxonomy of the ARCS theory reflected in the scale” (p. 286). On top of that, it has confirmed that the conditional nature of the ARCS Model is reflected in the RIMMS. This means that the underlying motivational theory of the ARCS Model holds true and is reflected in the RIMMS in our instructional setting outside the traditional classroom, namely a self-directed instructional setting aimed at working with technology.

Statistical arguments for preferring the RIMMS over the IMMS

The validation study described in this article showed that statistically, the RIMMS is preferred over the IMMS in a self-directed instructional setting with seniors. Looking at the content, form, and structure of the IMMS versus the RIMMS reveals a similar theoretical preference. Firstly, none of the RIMMS items are reverse items. Reverse items are commonly used to detect response biases like automatic answering strategies. Hinkin (1998) suggests keeping a measure short to minimize response biases caused by boredom or fatigue. Since the RIMMS consists of only 12 items compared to the original 36 items, the risk of such response biases is considered minimal.

Secondly, the representation of constructs by items is uneven in the IMMS, measuring Attention by 12 items, both Relevance and Confidence by 9 items, and Satisfaction by merely 6 items. Keller (2010) explains that “the primary reasons for the disproportionate numbers of items in the Attention and Satisfaction subscales are that boredom and lack of stimulation are such ubiquitous characteristics in instructional writing and the satisfaction category does not have as many points of connection to printed material as the others” (p. 282). His statement assumes that nine items per construct is norm, when this study showed that the RIMMS measures the four constructs well with three items each. The three remaining RIMMS items per construct might be conceptually different than the original six, nine or twelve items. It would be interesting to test for construct validity

comprehensively in a succeeding study, and especially for differences between the IMMS and the RIMMS. What is known now, is that the three remaining items per construct are distinctive, and that all three items measure the same parent construct. We named these parent constructs after the ARCS Model they represent, but additional construct validation studies—focusing on consequential, content, substantive, structural, external and generalizability issues (cf. Messick, 1995)—should reveal more about each of the twelve items. Also, in our study, the IMMS was used as a starting point. This resulted in the RIMMS, which consists of twelve IMMS items. Our results suggest that the RIMMS is a better measure of motivation in our setting, but additional construct validation studies should reveal if it is also the best self-report measure for motivation and its four constructs in our setting.

The four constructs and the conditional nature of the ARCS Model are reflected in the 12-item RIMMS. In other words, the RIMMS, with a greatly reduced number of items, measures the four ARCS constructs more parsimoniously than the original IMMS. Hinkin (1998) refers to Thurstone (1947), pointing out that “scales should possess simple structure, or parsimony. Not only should any one measure have the simplest possible factor constitution, but any scale should require the contribution of a minimum number of items that adequately tap the domain of interest” (p. 109).

Substantive arguments for preferring the RIMMS over the IMMS

The creation of the RIMMS was not solely based on statistical grounds: Our thorough (“A to Z”) validation approach encompassed deleting items based on statistical arguments, but only when this could be backed up by substantive arguments. This method allowed us to discover three substantive reasons for removing items from the original 36-item survey.

First of all, we discovered that two of the IMMS items were fact items. Both fact items were dispelled in forming the RIMMS (09R02 — *There were stories, pictures, or examples that showed me how this telephone could be important to some people*, and 18R05 — *This instruction manual contains explanations or examples of how people use the telephone*).

These fact items seem almost identical to each other, and do not reflect the effect of instructions on people, but are merely a manipulation check. Also, Study 1 showed that many participants experienced doubts filling out these items. Item 09R02, for instance, was scored highly by many participants, explaining they could imagine some people finding it important to have stories, pictures, or examples.

Second of all, two items (05S01 — *Completing the exercises gave me a satisfying feeling of accomplishment*, and 32S05 — *It felt good to successfully complete the exercises*) were dispelled, because people's answers to these Satisfaction items are linked to their success in performing tasks or exercises (see Loorbach et al., 2007). Removing these items seems especially important when the survey is used as a pre-test tool serving as a motivational needs assessment prior to instruction, and as such, prior to completing exercises.

A third reason for dispelling items in forming the RIMMS was ambiguity. We classified two items as negatively stated cause-effect items, stating something negative had happened because of a specific reason (ie. 12A04 — *This instruction manual is so abstract that it was hard to keep my attention on it*, and 26R07 — *This instruction manual was not relevant to me, because I already knew most of the content*). In these cases, half of the item could or could not be true for participants, while at the same time, the other half also could or could not be true. So it may or may not have been difficult for participants to keep their attention on the user instructions, because of its abstractness (or despite of it) or because of (an) entirely different reason(s), making it unclear to participants which answer category to choose. The two negatively stated cause-effect items were not included in the RIMMS.

Limitations and suggestions for future research

All in all, this study attempted to validate the IMMS from A to Z, and its results suggest using the RIMMS rather than the IMMS in self-directed instructional settings aimed at people likely to benefit from motivational instructions, to learn and perform tasks through self-study. However, the validation of this self-report measure of motivation is only in its infancy and our study had some limitations, that were mainly related to the relatively small sample size.

Much time and effort was put into gathering participants meeting our age and cell phone experience demands. On top of that, participants had to be willing to welcome a researcher in their own home and sit down for an hour and a half, using a cell phone and instructions, and filling out questionnaires. In total, 138 people participated. Even though this sample size proved sufficient to assess model fit, a greater sample size is recommended for future validation studies to ensure that statistical differences will be detected should they exist. In our studies, the validation study was part of a large usability study, resulting in relatively long sessions. We suggest limiting future validation sessions to a shorter time frame, to stimulate participation.

Another limitation of this study is that high RMSEA values may indicate serious model misspecifications. However, it is possible that the high RMSEA values in this study were influenced by the few degrees of freedom of the model. Chen, Curran, Bollen, Kirby, and Paxton (2008) suggest that to achieve a certain level of power or Type I error rate, the choice of cutoff values depends on model specifications, degrees of freedom, and sample size. The RMSEA is sensitive to the number of free parameters in the model. With few degrees of freedom in a model, one may expect higher RMSEA values (Loehlin, 2004).

Further validation studies should reveal whether the preference of the RIMMS over the IMMS holds true in self-directed instructional settings with other audiences. Furthermore, our study's self-directed instructional setting is comparable to more traditional situations in education, which could also benefit from the improved RIMMS. Therefore, we also suggest validating the RIMMS in the traditional instructional setting the ARCS Model and IMMS were originally designed for, like face-to-face classroom settings, as well as settings the ARCS Model has since been applied to, like self-directed computer-based instruction and distance education.

Finally, this study validated the IMMS as a post-test tool, measuring whether the four ARCS goals are met in instructional materials. The RIMMS seems to be a better post-test tool than the IMMS in our self-directed instructional setting. We suspect the RIMMS to also be a better pre-test tool than the IMMS in our setting, due to its parsimonious nature and the absence of fact items, ambiguous items, and items linked to experiencing success or failure, which in the case of pretesting, will not have taken place yet. An additional validation study should reveal whether the RIMMS is indeed also preferred for measuring motivation prior to applying ARCS strategies to instructional materials. If so, then this will contribute to determining people's motivational needs, and ultimately, to the quality of ARCS-based motivational instructions.

Acknowledgements

The authors would like to thank Nick Rushby, Editor of the journal, for his helpful guidance in the revision of this paper. We would also like to thank two anonymous reviewers for their thorough review of our manuscript and for their helpful comments.

References

- Anderson, J.C., & Gerbing, D.W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*, 411—423.
- Astleitner, H., & Hufnagl, M. (2003). The Effects of Situation-Outcome-Expectancies and of ARCS-Strategies on Self-Regulated Learning with Web-Lectures. *Journal of educational multimedia and hypermedia*, *12*(4), 361—376.
- Bellon, T., & Oates, R. (2002). *Best Practices in Cyberspace: Motivating the Online Learner* [WWW document]. URL <http://168.144.129.112/articles/best%20practices%20in%20cyberspace.pdf>
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (136—162). Thousand Oaks, CA: Sage.
- Chang, M., & Lehman, J.D. (2002). Learning Foreign Language through an Interactive Multimedia Program: An Experimental Study on the Effects of the Relevance Component of the ARCS Model. *CALICO Journal* *20*(1), 81—98.
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An Empirical Evaluation of the Use of Fixed Cutoff Points in RMSEA Test Statistic in Structural Equation Models. *Sociological Methods & Research* *36*(4), 462—494.
- Chisnell, D., & J. C. Redish (2005). Who Is the "Older Adult" in Your Audience? *Intercom* *52*(1), 10—14.
- Choi, H.J., & Johnson, S.D. (2005). The Effect of Context-Based Video Instruction on Learning and Motivation in Online Courses. *The American Journal of Distance Education*, *19*(4), 215—227.
- Chyung S. Y., Winiecki, D., & Fenner, J. A. (1999). Evaluation of effective interventions to solve the drop-out problem in adult distance education. In B. Collis & R. Oliver (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications* (51—55). Charlottesville, VA: AACE.
- Fan, X., Tompson, B., & Wang, L. (1999). Effects of sample size, estimation method, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling*, *6*, 56—83.
- Gerbing, D.W., & Anderson, J.C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, *11*, 572—580.
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*, 104—121.

- Holbert, R. L., & Stephenson, M. T. (2002), Structural equation modeling in the communication sciences, 1995—2000. *Human Communication Research*, 28(4), 531—551.
- Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: Comments, issues, and applications*. Thousand Oaks, CA: Sage, 158—176.
- Hu, L.-T., & 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.
- Huang, W., Huang, W., Diefes-Dux, H., & Imbrie, P.K. (2006). A preliminary validation of Attention, Relevance, Confidence and Satisfaction model-based Instructional Material Motivational Survey in a computer-based tutorial setting. *British Journal of Educational Technology*, 37(2), 243—259.
- Keller, J.M. (1983). Motivational Design of Instruction. In C.M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum Associates, 386—434.
- Keller, J.M. (1987a). Development and Use of the ARCS Model of Motivational Design. *Journal of Instructional Development*, 10(3), 2—10.
- Keller, J.M. (1987b). Strategies for Stimulating the Motivation To Learn. *Performance & Instruction* 26(8), 1—7.
- Keller, J.M. (1987c). The Systematic Process of Motivational Design. *Performance & Instruction*, 26(9), 1—8.
- Keller, J.M. (1999). Using the ARCS Motivational Process in Computer-Based Instruction and Distance Education. *New Directions for Teaching and Learning* 78, 39—47.
- Keller, J.M. (2010). *Motivational Design for Learning and Performance: The ARCS Model Approach*. New York, NY: Springer.
- Keller, J.M., & Kopp, T.W. (1987). An Application of the ARCS Model of Motivational Design. In C.M. Reigeluth (Ed.), *Instructional theories in action. Lessons illustrating selected theories and models*. Hillsdale, NJ: Lawrence Erlbaum Associates, 289—320.
- Loehlin, J. C. (2004). *Latent variable Models: An Introduction to factor, Path, and Structural Equation Analysis*. Mahwah, NJ: Lawrence Erlbaum.
- Loorbach, N., Karreman, J., & Steehouder, M. (2007). Adding Motivational Elements to an Instruction Manual for Seniors: Effects on Usability and Motivation. *Technical Communication*, 54(3), 343—358.

- Loorbach, N., Karreman, J., & Steehouder, M. (2013a). Confidence-Increasing Elements in User Instructions: Seniors' Reactions to Verification Steps and Personal Stories. *Technical Communication*, 60(3), 190—204.
- Loorbach, N., Karreman, J., & Steehouder, M. (2013b). Verification steps and personal stories in an instruction manual for seniors: Effects on confidence, motivation, and usability. *IEEE Transactions on Professional Communication*, accepted for publication, 2013.
- Means, T.B., Jonassen, D.H., & Dwyer, F.M. (1997). Enhancing Relevance: Embedded ARCS Strategies vs. Purpose. *Educational Technology Research and Development*, 45(1), 5—18.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5—8.
- Schwender, C., & Köhler, C. (2006). Introducing Seniors to New Media Technology: New Ways of Thinking for a New Target Group. *Technical Communication*, 53(4), 464—470.
- Shellnut, B., Knowlton, A., & Savage, T. (1999). Applying the ARCS Model to the Design and Development of Computer-Based Modules for Manufacturing Engineering Courses. *ETR&D*, 47(2), 100—110.
- Song, S.H., & Keller, J.M. (2001). Effectiveness of Motivationally Adaptive Computer-Assisted Instruction on the Dynamic Aspects of Motivation. *ETR&D*, 49(2), 5—22.
- Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (2001). *Using Multivariate Statistics* (4th edn.). Needham Heights, MA: Allyn & Bacon, 653—771.

Appendix A

The Instructional Materials Motivation Survey

(Keller, 2010, pp. 283—284)

Instructions

Instructional Materials Motivation Survey

There are 36 statements in this questionnaire. Please think about each statement in relation to the instructional materials you have just studied and indicate how true it is. Give the answer that truly applies to you, and not what you would like to be true, or what you think others want to hear. Think about each statement by itself and indicate how true it is. Do not be influenced by your answers to other statements.

Record your responses on the answer sheet that is provided and follow any additional instructions that may be provided in regard to the answer sheet that is being used with this survey. Thank you. Use the following values to indicate your response to each item.

1 (or A) = Not true

2 (or B) = Slightly true

3 (or C) = Moderately true

4 (or D) = Mostly true

5 (or E) = Very true

01C01 When I first looked at this lesson, I had the impression that it would be easy for me.

02A01 There was something interesting at the beginning of this lesson that got my attention.

03C02 This material was more difficult to understand than I would like for it to be.*

04C03 After reading the introductory information, I felt confident that I knew what I was supposed to learn from this lesson.

05S01 Completing the exercises in this lesson gave me a satisfying feeling of accomplishment.

06R01 It is clear to me how the content of this material is related to things I already know.

07C04 Many of the pages had so much information that it was hard to pick out and remember the important points.*

08A02 These materials are eye-catching.

09R02 There were stories, pictures, or examples that showed me how this material could be important to some people.

10R03 Completing this lesson successfully was important to me.

11A03 The quality of the writing helped to hold my attention.

12A04 This lesson is so abstract that it was hard to keep my attention on it.*

13C05 As I worked on this lesson, I was confident that I could learn the content.

14S02 I enjoyed this lesson so much that I would like to know more about this topic.

15A05 The pages of this lesson look dry and unappealing.*

16R04 The content of this material is relevant to my interests.

17A06 The way the information is arranged on the pages helped keep my attention.

18R05 There are explanations or examples of how people use the knowledge in this lesson.

- 19C06 The exercises in this lesson were too difficult.*
- 20A07 This lesson has things that stimulated my curiosity.
- 21S03 I really enjoyed studying this lesson.
- 22A08 The amount of repetition in this lesson caused me to get bored sometimes.*
- 23R06 The content and style of writing in this lesson convey the impression that its content is worth knowing.
- 24A09 I learned some things that were surprising or unexpected.
- 25C07 After working on this lesson for a while, I was confident that I would be able to pass a test on it.
- 26R07 This lesson was not relevant to my needs because I already knew most of it.*
- 27S04 The wording of feedback after the exercises, or of other comments in this lesson, helped me feel rewarded for my effort.
- 28A10 The variety of reading passages, exercises, illustrations, etc., helped keep my attention on the lesson.
- 29A11 The style of writing is boring.*
- 30R08 I could relate the content of this lesson to things I have seen, done, or thought about in my own life.
- 31A12 There are so many words on each page that it is irritating.*
- 32S05 It felt good to successfully complete this lesson.
- 33R09 The content of this lesson will be useful to me.
- 34C08 I could not really understand quite a bit of the material in this lesson.*
- 35C09 The good organization of the content helped me be confident that I would learn this material.
- 36S06 It was a pleasure to work on such a well-designed lesson.

* Asterisked items should be recoded prior to data analysis (1 = 5, 2 = 4, 4 = 2, and 5 = 1).

¹ 02A01 is the second item of the IMMS scale, and the first item of the A construct, 03C02 is the third item of the IMMS scale, and the second item of the C construct, etc. Codes were added for reference; in the original IMMS, items are numbered 1 through 36.

Figures and Tables

Figure 1: ARCS Model of Motivational Design

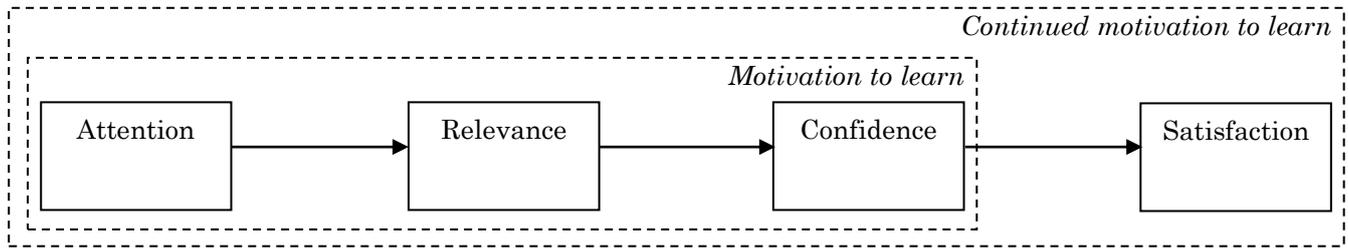


Figure 2: Path model in line with ARCS theory

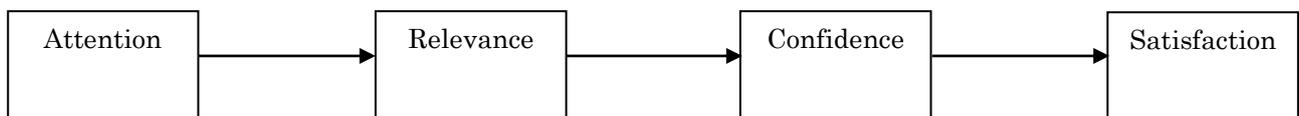
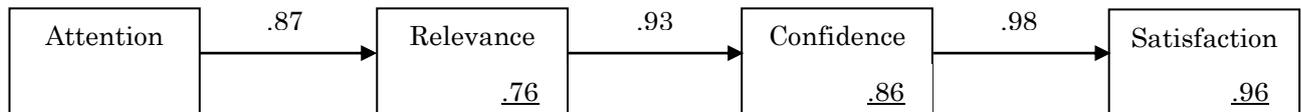


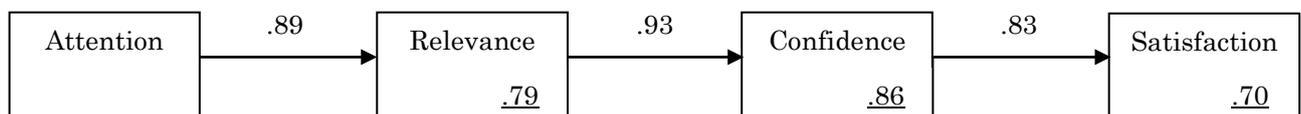
Figure 3: Standardized path coefficients of the ARCS Model of Motivational Design (Study 1, IMMS)*



* The outcomes are based on the IMMS study, but calculated using the remaining 12 IMMS items only (forming the RIMMS for Study 2).

Note. See Table 2 for the observed indicators of the latent constructs. Paths are significant at $p < .001$. Squared multiple correlations are underlined.

Figure 4: Standardized path coefficients of the ARCS Model of Motivational Design (Study 2, RIMMS)



Note. Paths are significant at $p < .001$. Squared multiple correlations are underlined.

Table 1. Correlation matrix of the Reduced Instructional Materials Motivation Survey (RIMMS) items, Study 1

		1	2	3	4	5	6	7	8	9	10	11	12
1	11A03	-	.46	.53	.37	.42	.53	.42	.42	.38	.48	.35	.35
2	17A06		-	.44	.33	.52	.43	.33	<i>.16</i>	.43	.60	.45	.49
3	28A10			-	.23	.45	.45	.38	.38	.41	.44	.47	.40
4	06R01				-	.28	.43	.32	.43	.38	.38	.24	.31
5	23R06					-	.53	.39	.33	.56	.52	.49	.59
6	33R09						-	.48	.39	.63	.48	.54	.62
7	13C05							-	.40	.56	.50	.61	.50
8	25C07								-	.47	.35	.26	.36
9	35C09									-	.50	.61	.76
10	14S02										-	.56	.59
11	21S03											-	.64
12	36S06												-

Note. Correlations significant at $p < .05$, non-significant correlation is in italic. 11A03 is the 11th item of the original 36-item IMMS scale, and the 3rd item of the original A construct.

Table 2. Descriptive statistics, factor loadings, squared multiple correlations, and Cronbach's alpha of the observed indicators to explain motivation

	Study 1				Study 2	
	M	SD	β	R^2	M	SD
Attention ($\alpha = .73$, Study 1; $\alpha = .90$, Study 2)					3.07	1.34
The quality of the text helped to hold my attention (11A03)*	2.83	1.13	.71	.50	3.02	1.55
The way the information is arranged on the pages helped keep my attention (17A06)*	2.69	1.12	.68	.46	3.12	1.46
The variety of reading passages, exercises, illustrations, etc., helped keep my attention on the user instructions (28A10)*	2.47	1.19	.69	.48	3.08	1.39
There was something interesting at the beginning of these user instructions that got my attention (02A01)	1.49	.84				
These user instructions are eye-catching (08A02)	2.64	1.27				
These user instructions are so abstract that it was hard to keep my attention on them (12A04)	3.51	1.38				
The pages of these user instructions look dry and unappealing (15A05)	4.18	1.04				
These user instructions have things that stimulated my curiosity (20A07)	2.89	1.24				
The amount of repetition in these user instructions caused me to get bored sometimes (22A08)	4.28	1.03				
I learned some things that were surprising or unexpected (24A09)	2.49	1.33				
These user instructions' style of writing is boring (29A11)	4.09	1.12				
There are so many words on each page that it is irritating (31A12)	3.67	1.41				
Relevance ($\alpha = .68$, Study 1; $\alpha = .82$, Study 2)				.76	3.38	1.30
It is clear to me how the content of these user instructions is related to things I already know (06R01)*	2.17	1.20	.49	.24	3.32	1.58
The content and style of writing in these user instructions convey the impression that being able to work with the telephone is worth it (23R06)*	3.03	1.34	.72	.52	3.34	1.45
The content of these user instructions will be useful to me (33R09)*	3.23	1.50	.79	.62	3.49	1.51
There were stories, pictures, or examples that showed me how this telephone could be important to some people (09R02)	2.62	1.28				
Completing the exercises successfully was important to me (10R03)	3.63	1.35				
The content of these user instructions is relevant to me (16R04)	2.96	1.45				
These user instructions contain explanations or examples of how people use the telephone (18R05)	2.82	1.43				
These user instructions were not relevant to me, because I already knew most of the content (26R07)	4.83	.54				
I could relate the content of these user instructions to things I have seen, done, or thought about before (30R08)	2.29	1.20				
Confidence ($\alpha = .73$, Study 1; $\alpha = .89$, Study 2)				.86	3.28	1.33
As I worked with these user instructions, I was confident that I could learn how to work well with the telephone (13C05)*	2.82	1.24	.68	.46	3.17	1.49
After working with these user instructions for a while, I was confident that I would be able to complete exercises with the telephone (25C07)*	2.39	1.24	.51	.26	3.41	1.51
The good organization of the content helped me be confident that I would learn to work with the telephone (35C09)*	2.72	1.27	.84	.70	3.27	1.44
When I first looked in these user instructions, I had the impression that these would be easy to work with (01C01)	2.26	1.09				
These user instructions were more difficult to understand than I would like for them to be (03C02)	2.74	1.39				
After having looked in these user instructions briefly, I felt confident that I knew what would be discussed in these instructions (04C03)	2.47	1.20				

Many of the pages had so much information that it was hard to pick out and remember the important points (07C04)	2.80	1.39				
The exercises with these user instructions were too difficult (19C06)	3.02	1.38				
I could not really understand quite a bit of the information in these user instructions (34C08)	3.21	1.30				
Satisfaction ($\alpha = .82$, Study 1; $\alpha = .85$, Study 2)				.96	2.79	1.26
I enjoyed working with these user instructions so much that I was stimulated to keep on working (14S02)*	2.27	1.23	.71	.50	2.37	1.47
I really enjoyed working with these user instructions (21S03)*	2.63	1.23	.76	.58	2.98	1.40
It was a pleasure to work with such well-designed user instructions (36S06)*	2.42	1.21	.85	.72	3.02	1.43
Completing the exercises gave me a satisfying feeling of accomplishment (05S01)	2.26	1.36				
The comments in these user instructions helped me feel rewarded for my effort (27S04)	2.08	1.09				
It felt good to successfully complete the exercises (32S05)	2.96	1.44				

* RIMMS items. Cronbach's alphas and overall means were calculated using these items only.

Note. The R^2 of a latent dependent predictor is the percentage of the variance in the latent dependent variable accounted for by the latent independent variable. The R^2 of an observed indicator is the estimated percent variance explained in that variable.

Table 3: Correlation matrix of the Reduced Instructional Materials Motivation Survey (RIMMS), Study 2

		1	2	3	4
1	A	-	.89	.87	.81
2	R		-	.93	.82
3	C			-	.83
4	S				-

Note. Correlations significant at $p < .01$. A = Attention, R = Relevance, C = Confidence, S = Satisfaction.