An extended short scale for measuring intrinsic motivation when engaged in inquiry-based learning

Manuela Heindl
Technical University Dresden, Dresden, Germany

The aim of this paper was to extend the German multidimensional measurement (KIM) to assess participants’ subjective experience related to a target activity. This device consists of: one dimension with three items to test experience-based hypothesising and four dimensions to test intrinsic motivation. The data was gathered in Austria in 2017/18 in various primary schools (age=10 years, n= 677) after an inquiry-based learning lesson. The EFA revealed 15 items were loading on five factors. The reliability (Cronbach’s α = .813 and Guttmann split-half coefficient = .834) and the validity (χ²(80) = 123.997; RMSEA = .0029 (90% CI = .018-.0038); CFI = .986; TLI = .982; SRMR = .029) proved significantly a good model fit. The six dimensional questionnaire EKIM can be used to test intrinsic motivation and when engaged in inquiry-based learning in primary schools.

Keywords: Intrinsic motivation; Short scale; Inquiry-based learning; Multidimensional measurement

Article History: Submitted 2 September 2018; Revised 24 December 2019; Published online 13 January 2020

1. Introduction

The term intrinsic motivation, according to Ryan and Deci (2000), is an important construct within schools and specifically refers to a human propensity to learn and assimilate. Derived from Latin ‘movere’, the term ‘motivation’ implies being moved to begin an activity. Therefore, a motivated person is characterised by being energised and feeling full of inspiration. People have different levels and different kinds of motivation. The nature and the focus of the motivation can vary. In Self-Determination Theory, Deci and Ryan (1985) discuss different types of motivation based on different goals, the most basic one being intrinsic motivation. The term Intrinsic Motivation defines carrying out an action because it is interesting or enjoyable. This concept is regarded as an important phenomena for teachers, as it is the source of natural learning and the achievement can be coordinated. It, therefore, results in high-quality learning and creativity. This assumption can be also verified by several recent studies. Froiland and Worrell (2016) examined the relationship a higher academic performance and intrinsic motivation in 1575 students in an ethnically and diverse high school. A positive relationship can be found between a higher level of inquiry-based instruction and a higher achievement in students of science (Velooa, Perumalb, & Vikneswary,
Aarepattamnill, Freeman and Klinger (2011) revealed a study about positive predictive effects of intrinsic motivation on academic achievement for Indian immigrants and adolescents. Several studies can be found that show a positive correlation between academic achievement and intrinsic motivation (Saeed, & Zyngier, 2012; Corpus, McClintic-Gilbert, & Hayenga, 2009; Lee, McInerney, Liem, & Ortiga, 2010). Although, according to Garon-Carrier’s et al. (2015), the cross-lagged model showed that achievement predicted intrinsic motivation from grades 1 to 4 in Canada, however, intrinsic motivation did not predict achievement at any time. Although intrinsic motivation cannot be guaranteed within higher level achievement of students, it is a basic element of inquiry-based learning and indirectly effects student achievement. It is therefore important to maintain it in classroom. Biley and Smith (1999) talk about a decrease in student motivation when a more traditional teaching style of teaching is taken on by the teacher. According to Saunders-Stewart, Gyles and Shore (2012), the current trend in teaching goes towards a student-centred approach.

Within a more traditional learning setting teachers are seen as the source of information. They lead and show their power over students who are mainly receiving and storing knowledge. More recently A new teaching method has evolved which uses problem-solving in a research setting. The terms inquiry-based learning or enquiry-based learning, which is the equivalent in the UK, can be compared to problem-based learning. The latter focusses on a problem rather than a whole scenario. It must be stressed that both approaches have emerged in order to promote the future career of students (Biley & Smith, 1999). On the job, they need to be autonomous, reflective and know how to react to problems. Critical thinking is very essential in an ever-changing and complex environment. According to Reitinger (2016) students usually work in small groups on a problem in an inquiry-based learning setting and develop a research question with a plan to test their hypothesis in an experiment. Their data or outcomes are presented in class or to another audience in order to discuss and evaluate their findings. According to Deci and Ryan (1985) the experience of being competent, and autonomous is based on the concept of intrinsic motivation. Therefore, inquiry-based learning is closely linked to this concept. Furthermore, inquiry-based learning, is, according to Reitinger (2016), driven by an interest in discovery. Self-determined inquiry is therefore linked to the interest of the learners. This disposition is a criterion for inquiry learning settings and is visible as a dimension within the CILI, the Criteria of Inquiry Learning Inventory by Reitinger (2016). This device and other kinds of assessment instruments were developed in order to measure interest in an activity. However, researchers who want to capture the level of interest in primary school pupils are faced with the problem of finding a short, reliable and valid assessment instrument that is suitable for this age group.

In Germany the short scale assessment instrument of Wilde et al. (2009) has been established to assess teenager’s level of interest as it had been tested for its psychometric quality. It is a German translation of the English standardized Intrinsic Motivation Inventory (IMI) questionnaire, which has been used extensively over the decades. It is based on Deci’s and Ryan’s Self-Determination Theory (1985). The IMI is a multi-dimensional measurement tool to investigate intrinsic motivation. To measure the seven-dimensional construct 45 items are needed. Assuming an average processing time of four minutes per item (Schwarzer & Jerusalem, 2002) for adults, this is too time consuming. However, answering the questions in school might take even longer, because young learners still have to focus on reading and understanding the construct first. The duration of the survey plays an important role in ensuring that a research project is approved by the competent school board or government. It’s also in the researcher’s interest to design short questionnaires using familiar words to ensure reliability because of its risk that a questionnaire is aborted or insufficiently filled in when fatigue or frustration sets in. The original scale of IMI was already extended by Deci and Ryan (1985) by two dimensions, but the German version is not yet extended. The short German questionnaire ‘Kurzskala intrinsischer Motivation’ (KIM) by Wilde et al. (2009) allows researchers to examine the level of interest in secondary school students, hence it has to be tested if it is also suitable for elementary school pupils. The KIM as an assessment
instrument to measure the level of interest is sufficiently reliable and has proven to be theory-compliant. The goal of this research is to develop the EKIM short for ‘Erweiterte Kurzskala intrinsischer Motivation’ - a questionnaire in German to test the level of interest in inquiry-based learning settings. Based on the scale constructed by Deci and Ryan (1985), a short and reliable one-dimensional scale for measuring value/usability is translated and validated within the framework of the construct "intrinsische motivation". It will be attached to the existing four-dimensional scale of Wilde et al. (2009), which will be tested if it is suitable for learners in elementary schools. As this project is carried out in an inquiry-based lesson, another dimension will be added to the EKIM. Equally, Reitinger’s (2016) ‘experience-based hypothesizing’ dimension for young learners should be translated into German and validated. The short scale is intended to explain the level of interest when engaged in inquiry-based learning.

1.1. Research Question

As the validitated instrument IMI can test the level of intrinsic motivation, it should be translated into German and also validated for elementary students:

Is the IMI translated into German, which tests intrinsic motivation, valid for elementary school pupils in inquiry-based learning settings?

2. Method

2.1. Participants and Data Collection

As part of the construction and validation of the scale, two samples were taken. The characteristics are shown in Table 1. Wave 1 (N = 207) was carried out half a year later than Wave 2 (N = 470) in elementary school classes. The data collection and the project was the same for both waves: After a two-hour inquiry-based learning unit the data was collected in form of a paper questionnaire. The assignment of the children to wave 1 or 2 was random. It took the elementary school pupils 20 minutes to fill in the EKIM questionnaire.

Table 1.
Descriptives for each time frame

<table>
<thead>
<tr>
<th>Time frame</th>
<th>Age</th>
<th>N</th>
<th>Boys</th>
<th>Girls</th>
<th>Mean</th>
<th>Variance</th>
<th>Std. Deviation</th>
<th>Nr. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017 wave 1</td>
<td>10y</td>
<td>207</td>
<td>102</td>
<td>105</td>
<td>8.1300</td>
<td>66.847</td>
<td>2.91694</td>
<td>15</td>
</tr>
<tr>
<td>2018 wave 2</td>
<td>10y</td>
<td>470</td>
<td>237</td>
<td>233</td>
<td>1.1958</td>
<td>46.457</td>
<td>2.10195</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>10y</td>
<td>677</td>
<td>339</td>
<td>338</td>
<td>1.1958</td>
<td>46.457</td>
<td>2.1019</td>
<td>15</td>
</tr>
</tbody>
</table>

2.2. Instrument Development

The extended short scale or EKIM for intrinsic motivation was the aim of this research and consists of five subscales, each containing three items. The answers of very item range from (1) "not true" to (5) "true". The items were translated from English into German by a team of educators, a director and the researcher. This procedure was carried out according to the recommendations of the International Test Commission (ITC, 2018). It was then translated back into German to see if the translation was successful. This version was reviewed by an Irish educator and compared to the original. There were hardly any differences and the changes were discussed and slightly modified. The aim of this modification was to increase the comprehensibility of the items for young learners. The sentence structure was simplified and adapted to the existing questionnaires, so that almost all items have the same sentence starter. In addition, the answer scale was reduced to 5 instead of having 7 possible answers as suggested and applied by the University of Rochester (2016). The modified items were then tested in a pilot study (N = 21, age = 10 years, Mean = 53.9048, SD = 18.2808, Cronbach’s α = 0.979). The goal was to test the generated items. One key issue from the pilot study was that a word in the questionnaire was too abstract and allowed different interpretations (valuable – „wertvoll“) so that it was changed (into useful – „brauchbar“). In addition, an discussion with the pupils and the teacher was sought to discuss the layout and
simplicity of the questionnaire. A 5-step response scale was chosen, with the following category names accompanied by smileys: 1) "strongly disagree" 2) "slightly disagree" 3) "agree slightly " 4) "agree" 5) "strongly agree" (Rohrmann, 1978). The items and response scale are listed in the appendix. Only demographic questions such as school, school class, gender and mother tongue were included in the survey. In the preparation phase of the project, the submission of the project application was approved by the Department of Education. Then a written consent to participate in the survey was sent to the educational board of the primary school and later on to the legal guardians of the children of the respective classes.

3. Findings

2.2. Objectivity

The objectivity refers to different stages of the implementation and is the degree to which a measurement is independent from the examiner in the performance, evaluation and interpretation. The survey was conducted by independent researchers working at different teacher training colleges. The analysis was defined in advance according to a numerical and categorical answer option of the questionnaire. The values of the items are clearly defined and left no room for any misinterpretation. The interpretation objectivity can also be considered as fulfilled by standardizing the analysis and assigning a measured value to the answers.

2.2. Reliability

Reliability is the measurement accuracy of a scale, and three methods were calculated for the used scale: split-half reliability, retest reliability and scale homogeneity. The latter suggests a very good stability with a value of $\alpha = .813$. The interpretation of the split-half and the retest reliability also show no concerns about the accuracy of the questionnaire. The results show that both test halves already have satisfactory reliabilities. Although the two test halves are not exactly parallel ($Part 1 = .648$, $Part 2 = .715$), the correlation between the test halves is satisfactory ($r = .728$). The Guttman split-half coefficient ($0.834$) indicates a slightly higher value than the calculated Cronbach $\alpha$ ($0.813$). Overall, the reability coefficients lie in a range that is rated as good.

2.3. Validity

Content validity reflects the precise wording of the items for the construct to be measured. The construct was defined a priori sufficiently and the items were judged by a panel of experts regarding their validity. A factorial validity was tested on the basis of sample 1. A confirmatory factor analysis was calculated in Lavaan version 0.5-23 (Rosseei, 2017) in R version 0.97.551 (R Core Team, 2018) and confirms the five-factorial structure of the scale in wave 1 ($\chi^2 (80) = 157.655, p < .001; \text{RMSEA} = .06 (90\% CI = .053-.084); \text{CFI} = .958; \text{TLI} = .945; \text{SRMR} = .048$) and wave 2 ($\chi^2 (80) = 130.434, p < .001; \text{RMSEA} = .037 (90\% CI = .025-.048), \text{CFI} = .972, \text{TLI} = .964, \text{SRMR} = .0034$). The five-factorial model showed significantly better results than a one-factorial solution ($\chi^2 (77) = 558.942; p < .001; \text{RMSEA} = .096 (90\% CI = .089-.0104); \text{CFI} = .832; \text{TLI} = .801; \text{SRMR} = .058$). The model with five latent and fifteen manifest variables is confirmed at both waves and for the whole sample ($\chi^2 (80) = 123.997, p < .001, \text{RMSEA} = .0029 (90\% CI = .018-.0038), \text{CFI} = .986; \text{TLI} = .982; \text{SRMR} = .029$). More details can be found in the results of the CFA.

2.4. Explorative Factorial Analysis (EFA)

Before a confirmatory factor analysis can be calculated, an exploratory factor analysis should confirm the structure of the questionnaire (Schermelleh-Engel, Moosbrugger, & Mueller, 2003). The KMO and Bartlett’s test are satisfactory with a result of $0.901 (\chi^2 (105) = 3237.665, p < 0.001)$, so that a principal axis analysis can be performed. This method was chosen because it requires fewer prerequisites and is therefore more robust (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Varimax was chosen as the rotation method because this method reduces small factor loadings and thus simplifies the interpretation of the factors. Kaiser normalization considers all variables as
equally important for rotation (Gorsuch, 1983). All variables load with more than 0.3 on the corresponding factor, with the exception of one item (ssat). There are several reasons for this, but the variable does not need to be further explored because the factor „competence“ has already been verified with a CFA by Wilde et al. (2009).

2.5. Confirmatory Factorial Analysis (CFA)

The factorial structure of the short scale was verified by a confirmatory factor analysis, a CFA (Kline, 2013). The following indices were used to show the quality criteria for the model fit (Beauducel & Wittmann, 2005): χ² (df, p), Comparative Fit Index (CFI, satisfactory fit > .95), Tucker Lewis Index (TLI, satisfactory fit > .95), root mean square error of approximation (RMSEA; satisfactory fit < .08). Table 2 shows the input in R by with the help of the add-on package called Lavaan. On the left side of the symbol = ~ (measured by) is a latent variable that is measured or operationalised by the three observed variables to the right. As the data contains missing values, these cases would usually be excluded from the analysis but in the ML estimation, it is possible to include this data to extract more information. Therefore, the DWLS method was preferred and this method gave no better results than the ML, which shows a reliable result.

Table 2.
*Input in R*

```r
> model <- '
+ inter =~ Ijoy + Iinter + Ifunny
+ skill =~ Sskill + Ssat + Sgood
+ choice =~ Cchoice + Cmy + Cguide
+ value =~ Vessent + Vuse + Vimp
+ hypo =~ Hassump + Hthink + Hanswer
```

Table 3 shows a list of the parameters tested and their significant positive loading on the factor. In the first column are the non-standardised parameter estimates followed by the following values: the standard error estimate, the standard deviation (SD), z value, and the probability p value for each calculated parameter.

Table 3.
*Latent Variables*

|                   | Estimate | SE  | z     | P(>|z|) |
|-------------------|----------|-----|-------|---------|
| inter =~          |          |     |       |         |
| Ijoy              | 1.000    |     | 9.462 | 0.000   |
| Iinter            | 0.935    | 0.099|       |         |
| Ifunny            | 0.814    | 0.090| 9.036 | 0.000   |
| skill =~          |          |     |       |         |
| Sskill            | 1.000    |     | 5.150 | 0.000   |
| Ssat              | 0.813    | 0.158|       |         |
| Sgood             | 1.036    | 0.071| 14.550| 0.000   |
| choice =~         |          |     |       |         |
| Cchoice           | 1.034    | 0.073| 14.216| 0.000   |
| Cmy               | 1.220    | 0.114| 10.677| 0.000   |
| Cguide            |          |     |       |         |
| value =~          |          |     |       |         |
| Vessent           | 1.000    |     | 21.746| 0.000   |
| Vuse              | 0.995    | 0.046|       |         |
| Vimp              | 0.709    | 0.041| 17.410| 0.000   |
| hypo =~           |          |     |       |         |
| Hassump           | 1.000    |     | 13.379| 0.000   |
| Hthink            | 1.017    | 0.076|       |         |
| Hanswer           | 1.227    | 0.086| 14.348| 0.000   |
Table 4 shows the variances for the 15 items in the questionnaire, which should not exceed the value of 1. They explain the variance in this variable in relation to the latent construct.

Table 4.

| Variance | Estimate | SE  | z     | \(P(>|z|)|
|----------|----------|-----|-------|------------------|
| 1.Ijoy   | 3.944    | 0.225 | 17.521 | 0.000            |
| 2.Inter  | 0.399    | 0.038 | 10.582 | 0.000            |
| 3.Ifunny | 0.658    | 0.044 | 15.088 | 0.000            |
| 4.Skill  | 0.608    | 0.050 | 12.169 | 0.000            |
| 5.Ssat   | 7.932    | 0.441 | 18.006 | 0.000            |
| 6.Sgood  | 0.529    | 0.049 | 10.741 | 0.000            |
| 7.Cchoice| 0.830    | 0.071 | 11.721 | 0.000            |
| 8.Cmy    | 0.853    | 0.074 | 11.470 | 0.000            |
| 9.Cguide | 4.394    | 0.266 | 16.545 | 0.000            |
| 10.Vessent| 0.500   | 0.044 | 11.373 | 0.000            |
| 11.Vuse  | 0.461    | 0.042 | 10.895 | 0.000            |
| 12.Vimp  | 0.669    | 0.042 | 15.839 | 0.000            |
| 13.Hassump| 0.787   | 0.051 | 15.338 | 0.000            |
| 14.Hthink| 0.648    | 0.045 | 14.546 | 0.000            |
| 15.Hanswer| 0.574  | 0.048 | 12.008 | 0.000            |
| inter    | 0.755    | 0.154 | 4.907  | 0.000            |
| skill    | 0.655    | 0.071 | 9.215  | 0.000            |
| choice   | 0.938    | 0.101 | 9.266  | 0.000            |
| value    | 1.067    | 0.088 | 12.095 | 0.000            |
| hypo     | 0.516    | 0.063 | 8.154  | 0.000            |

The results indicate a good model fit \(\chi^2 (80) = 123.997\) because the values of the RMSEA (.0029, 90\% CI = .018-.0038) and the SRMR (.029) are below .05 and the indices named CFI (.986) and TLI (.982) are close to 1. In summary, the results indicate that the scale with the fifteen items captures the theoretical constructs underlying this questionnaire. Thus it can be used in primary schools without any concerns.

4. Conclusion and Discussion

The aim of this work was to develop a German extended short scale or EKIM for intrinsic motivation in inquiry-based learning settings. For this reason, a pilot study was carried out and the items were optimised for the research phases in 2017 and 2018. The project was conducted by independent researchers from different teacher training colleges. 677 pupils in Austria at the age of 10 years old were asked to participate in an inquiry-based learning lesson in science and then rate their experience after the activity in a questionnaire. This device contains five dimensions to test intrinsic motivation and one dimension to test experience-based hypothesizing. The parameters for reliability (Cronbach’s \(\alpha = .813\) and Guttman split-half coefficient = 0.834) and the validity were calculated \(\chi^2 (80) = 123.997; \text{RMSEA} = .0029 (90\% \text{CI} = .018-.0038), \text{CFI} = .986, \text{TLI} = .982, \text{SRMR} = .029\) in order to check the psychometric quality of the translated scale and it proved a good model fit.

The findings of this study are congruent with other studies in which researchers used an intrinsic motivation inventory in order to assess the experience in inquiry-based learning settings. For example, Dickson (2010) examined the anxiety and motivation in nurses in an inquiry-based learning lesson. As the nurses needed to be autonomous practitioners it was evident to use an intrinsic motivation inventory. Nooijen (2018) has also translated the Intrinsic Motivation Inventory into Dutch although has not carried out a CFA or EFA as the aim was to examine the
intrinsic motivation in 55 students in an inquiry-based lesson. Reitinger (2016) has based his questionnaire on inquiry-based learning with the Intrinsic Motivation Inventory. Therefore, in inquiry-based learning lessons it is essential to use the IMI and the CILI. This paper aimed to combine the two questionnaires and translate them into German.

Further studies could include more dimensions on inquiry-based learning such as authentic exploration, critical discourse and conclusion-based transfer by (Reitinger, 2016). The inventory already exists in English and needs to be translated in German. Furthermore, it could be useful to measure secondary school students to expand the target group. Finally, the validity should be also tested with participants outside Austria.

The EKIM can be used in primary school after a lesson to verify the intrinsic motivation. The dimension experience-based hypothesizing is particularly suitable for inquiry-based learning lessons. The scale can be used in different school subjects. The use of the EKIM scale to measure intrinsic motivation and experience-based hypothesizing in primary schools can thus be recommended.

Acknowledgements

A sincere thank you to Abigail Brown for constructive criticism of the manuscript.

References


R Core Team. (2018). R. Vienna, Austria.
Appendix:

The questionnaire of the standardised assessment instrument from the Intrinsic Motivation Inventory (IMI) from Deci and Ryan (1985) was reduced to three items for: interest / pleasure, perceived competence, perceived freedom of choice and value / utility:

---------------------------
Answering scale:
1) "strongly disagree" 2) "slightly disagree" 3) "slightly agree" 4) "agree" 5) "strongly agree"

IMI or Intrinsic Motivation Inventory is the original version by Deci and Ryan to test intrinsic motivation. KIM or 'Kurzskala intrinsicher Motivation' stands for 'short scale Intrinsic motivation inventory' of Wilde et al. based on the IMI by Deci and Ryan. EKIM or Erweiterte Kurzskala intrinsischer Motivation stands for the extended short scale of the intrinsic motivation inventory for inquiry-based learning based on the translated version by Wilde et al. on the IMI by Deci and Ryan.

The term "in the exhibition/in der Ausstellung" represents a placeholder and can be replaced with another activity.

Interest / enjoyment---------------------------
Antwortskala:
1) „trifft nicht zu,” 2) „trifft wenig zu”, 3) „trifft etwas zu”, 4) „trifft ziemlich zu”, 5) „ trifft zu”

Interesse/Vergnügen (interest/enjoyment)
1. Die Tätigkeit in der Ausstellung hat mir Spaß gemacht.
2. Ich fand die Tätigkeit in der Ausstellung sehr interessant.
3. Die Tätigkeit in der Ausstellung war unterhaltsam.

Wahrgenommene Kompetenz (perceived competence)
6. Ich glaube, ich war bei der Tätigkeit in der Ausstellung ziemlich gut.

Wahrgenommene Wahlfreiheit (perceived choice)
7. Ich konnte die Tätigkeit in der Ausstellung selbst steuern.

Druck/Anspannung (pressure/tension)
12. Ich hatte Bedenken, ob ich die Tätigkeit in der Ausstellung gut hinbekomme.

-----------------
Kurzskala zum erfahrungsbasierten Hypothetisieren zum Forschenden Lernen (CILI) stands for short scale for hypothesising in inquiry-based lessons

13. Ich glaube, dass die Tätigkeit brauchbar für mich ist.
15. Ich glaube, dass die Tätigkeit wichtig ist.

EKIM - Erweiterte Kurzskala Intrinsischer Motivation (KIM) von Wilde et al. basierend auf IMI Deci und Ryan

Wert/Brauchbarkeit (value/usefulness)
13. Ich glaube, dass die Tätigkeit brauchbar für mich ist.
15. Ich glaube, dass die Tätigkeit wichtig ist.

Validierung: Ich würde gerne an einem weiteren (forschenden) Unterricht teilnehmen.