UNIVERSITY OF SZEGED
DOCTORAL SCHOOL OF EDUCATION

THE ASSESSMENT OF DOMAIN-GENERAL AND DOMAIN-SPECIFIC PROBLEM SOLVING IN INDONESIA: EXPLORING THE ROLE OF INDUCTIVE REASONING, SCIENTIFIC COMPETENCY AND AFFECTIVE FACTORS

PHD DISSERTATION SUMMARY
BY
AZIZUL GHOFAR CANDRA WICAKSONO

SUPERVISOR: DR. ERZSÉBET KOROM

SZEGED, HUNGARY
2024

1. Introduction

The digitalisation and technological development have raised concern among people about adapting to dynamic change in their social and daily lives. Those who can adapt to these changes will benefit their future work. However, individuals who lack adequate preparation may face insecurity and vulnerability in the workplace and social environment (OECD, 2017). To address this challenge, schools must prepare students with skills and abilities for the changes that are more rapid than before. Dealing with changes and unfamiliar situations is within the scope of the problem-solving construct (Greiff et al., 2012). Problem solving refers to the cognitive process of transforming a given situation into a solved situation when there is no obvious method or solution available (Mayer & Wittrock, 2006). It includes the willingness to engage in such a situation to understand and resolve the solutions (OECD, 2014). Problem solving has been studied for decades and is still considered an important skill in the 21st century (Baker & Mayer, 1999; Care et al., 2016; Csapó & Funke, 2017; Dindar et al., 2022; Wirth & Klieme, 2003).

The integration of problem solving in education has been applied in curricula, for example, the National Research Council (2012) specified problem solving as one of the core skills in K-12 science education. The Indonesian national curriculum also included problem solving as one of the core competencies to attain secondary education (BSNP, 2013). The integration of problem solving as a core component in curricula has led to a surge in research in the area of learning and instruction. However, in the area of assessment, There is less comprehensive information about students' problem-solving ability in Indonesia. Wicaksono & Korom (2022) conducted a review on problem solving assessment in Indonesia and found few standardised problem-solving assessment tests. Indeed, there is no recent study of comprehensive problem-solving assessment in the Indonesian context, which makes it important to conduct a study on domain-general and domain-specific problem-solving assessment.

The study of problem solving requires a dynamic connection from both cognitive and affective factors. Problem solving in a specific domain is strongly connected to prior knowledge to understand the problem situation, while general problem solving does not necessarily require prior knowledge, but is related to intelligence and reasoning process (Greiff & Neubert, 2014). Furthermore, affective factors such as attitude and motivation create a learning environment that enhances problem-solving activities. Thus, including these factors in the assessment process and analysing their connections to problem solving provides comprehensive information for improving problem solving.

2. Research questions and hypotheses

The present study aimed to measure students' problem-solving skills in Indonesia. Since the problem-solving measurement was categorized into two domains, general and specific domains, the measurement of problem-solving was also conducted in two main tests performance, domain-general problem-solving test, and domain-specific problem-solving test which presented in a science context. The other aim of this study was to investigate factors influencing problem-solving performance from the cognitive (inductive reasoning and scientific competency) and affective (attitude towards science and science motivation) sides. The research questions of the study were:

RQ 1: What are the students' domain-general and domain-specific problem solving, inductive reasoning and scientific competence, as well as their attitudes and motivation towards science?

RQ 2: Is there a significant difference in students' domain-general and domain-specific problem solving, inductive reasoning, and scientific competence, as well as their attitudes and motivation towards science based on their gender?

RQ 3: To what extent do cognitive and affective factors, gender, and socioeconomic status influence problem-solving ability?

To answer this question, hypotheses are formulated based on the theoretical background. Thus, the hypothesis for the research questions is described below.

H1: There is a significant difference between genders in domain-general problem-solving performance, with males outperforming females (Greiff et al., 2018; Wüstenberg et al., 2014).

H2: There is a significant difference between genders in domain-specific problem-solving performance, with males outperforming females (Gok, 2014; Soto-Ardila et al., 2022).

H3: There is a significant difference between genders in inductive reasoning performance (Blum et al., 2016), when males performed better in figural and numerical tasks (Waschl & Burns, 2020).

H4: There is a significant difference between genders in scientific competency, when females outperform males (OECD, 2004; Reilly, 2012).

H5: There is a significant difference between genders in attitude towards science, when females outperform males (Aini et al., 2019; Liu et al., 2010; Mihladiz et al., 2011).

H6: There is a significant difference between genders in science motivation, when females outperform males (Oppermann et al., 2021; Schürmann & Quaiser-Pohl, 2022).

H7: Inductive reasoning is associated with problem solving, meaning that students with higher inductive reasoning ability will perform better in domain-general problem-solving and domain-specific problem-solving tasks (Molnár et al., 2013).

H8: Scientific competency that represents domain knowledge is associated with problem solving (Glaser et al., 2009; Wang et al., 2013), meaning that students with higher scientific competency will show better performance in domain-general problem-solving and domain-specific problem-solving tasks.

H9: A positive attitude toward science will lead to higher performance in domain-general and domain-specific problem solving (Guven & Cabakcor, 2013).

H10: Higher motivation is associated with higher performance in domain-general and domain-specific problem solving (Bat Or, 2014).

H11: Gender difference is associated with domain-general and domain-specific problem-solving performance (Huang & Chen, 2016).

H12: Higher socio-economic status is associated with higher performance in domain-general and domain-specific problem solving (Chiu, 2022).

3. Methods

The present study was classified as a cross-sectional study with a descriptive approach to address the assessment of the problem-solving ability of students in Indonesia. An associational research design was also implemented to determine the relationship between the measured variables. The participants in this study were grade 10th secondary school students from Java, Indonesia. Clustered random sampling was performed by selecting several districts in Java and randomly selecting schools located in the urban area. A total of 1243 students participated in this study, 36.8% males and 63.2% females (M_{age} = 16.78 years). The instruments used in this study involved several tests and questionnaires, including:

a. MicroDYN test to measure domain-general problem solving. There are two phases in the MicroDYN test known as knowledge acquisition and knowledge application. The MicroDYN test in this study consists of 20 items and both phases have 10 items.
b. Science problem solving (SPS) test exhibits the problem-solving measurement in the specific domain. A total of 11 multiple-choice items were used to measure science problem-solving ability. Among the test items, 5 items represent the first category of science problem solving, *identifying the problem* (IP), and 6 items represent the second category, *generating solution* (GS).
c. Inductive reasoning test (IR), includes: figure series (FS), figure analogy (FA), number analogy (NA), and number series (NS) (Csapó, 1997; Pásztor et al., 2017). The test consists of 32 items, each part composed of 8 items.
d. Scientific competency test (SC), consists of two competencies, *explaining scientific phenomena* which is represented by 8 items, and *interpreting and evaluating scientific data or evidence* which contains 7 items (Wicaksono & Korom, 2023b).
e. Attitudes Towards Science Questionnaire (ATSQ) with a total of 27 items (Wicaksono & Korom, 2023a). It is divided into four latent variables: enjoyment (8 items), anxiety and difficulty (7 items), participation in science learning and activities (7 items), and the value of science (4 items).
f. Science Motivation Questionnaire II (SMQ II), with a total of 25 items (Glynn et al., 2011). It consists of five variables, including intrinsic motivation, career motivation, self-determination, self-efficacy, and grade motivation, with each variable having 5 items.

The background questionnaire was applied in this study to get additional information about students' gender and socioeconomic status (SES). The scoring type for the test was assigned as 1 point for the correct answer and the incorrect answer was assigned as 0 point. The scoring for the questionnaires follows the Likert scale ranging from 1 to 5 (1 = strongly disagree, 2 = disagree, 3 = neutral, rather agree or disagree, 4 = agree, and 5 = strongly agree). All tests and questionnaires were generated in English, and then back and forward translation was done into Indonesian language with the assistance of two reviewers. The instruments were administered via the online electronic diagnostic assessment system (eDia) (Csapó & Molnár, 2019).

Data undergoes a quantitative analysis to address research objectives. First, the validation of the instruments was presented with a focus on construct validity and reliability analysis. Confirmatory factor analysis (CFA) was used to ensure the construct of the tests based on the theoretical framework. We apply the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of the approximation (RMSEA), and the standardised root mean square residual (SRMR) as a parameter to describe the model with cutoff value as follows: CFI > 0.90; TLI > 0.90; RMSEA < 0.08; SRMR < 0.08 (Hu & Bentler, 1999). Furthermore, multidimensional Rasch model analysis was also performed to verify the psychometry properties and quality of the test items (Boone et al., 2014). The descriptive analysis was performed to profile the students' ability and their level in each variable measure. We used logit values for standardized measurement to describe students' ability levels. This logit value was generated from Rasch's analysis representing the level of student ability in each test and questionnaire responses based on standardization estimates (Boone, 2016). The comparison analysis was conducted using the t-test to compare the level of student ability according to their gender. To answer the connection between the cognitive and affective factors and problem-solving ability, we performed a structural equation model (SEM) with a maximum likelihood (ML) estimator to analyse the relationship between observed and latent variables (Khine, 2013).

4. Results

4.1. Instruments validation

The reliability analysis was performed to check the instruments internal consistency, resulting in acceptable result for MicroDYN test ($\alpha = 0.849$ and $\omega = 0.859$), science problem solving ($\alpha = 0.726$ and $\omega = 0.734$), inductive reasoning ($\alpha = 0.895$ and $\omega = 0.894$), scientific competency ($\alpha = 0.775$ and $\omega = 0.781$), attitude towards science ($\alpha = 0.834$ and $\omega = 0.827$), and science motivation ($\alpha = 0.949$ and $\omega = 0.949$). For construct validity, confirmatory factor analysis (CFA) was employed, resulting in an acceptable model fit for all tests and questionnaires, with CFI > 0.90; TLI > 0.90; RMSEA < 0.08; SRMR < 0.08, and factor loading higher than 0.4. Further analysis with the Rasch model was performed to measure the quality of the individual items based on the fitting index. We performed multidimensionality analysis to check the fit index of the items resulting in an acceptable fit for all instrument items within the range of infit MNSQ value between 0.5–1.5.

4.2. The profile of students' problem solving, and cognitive and affective factors

4.2.1. The profile of students' domain-general problem solving

The overall student performance has a mean logit value of -3.93 (SD = 1.80) which shows a lower level of ability than the average (below logit 0). In knowledge application, the mean logit value of the students reaches -3.77 (SD = 2.18). This result also applies to knowledge application when the average student score is -3.98 (SD = 0.46). The comparison between male and female students in domain-general problem-solving ability revealed a significant difference between male and female performance (supporting H1), where male students outperformed females ($M_{male} = -3.58$, $SD_{male} = 2.09$; $M_{female} = -4.13$, $SD_{female} = 1.58$; t = 5.18, $p = .000$).

4.2.2. The profile of students' domain-specific problem solving

The profile of the ability of the students to solve science problems showed a mean average of -0.88 logit (SD = 1.43). In the IP category, the students received average logit values of -0.93 (SD = 1.51), which is lower than the average score of the GS category which reached -0.77 logit (SD = 1.57). For gender comparison, we found no significant different difference between male and female students' scores in science problem solving ($M_{male} = -0.96$, $SD_{male} = 1.49$; $M_{female} = -0.84$, $SD_{female} = 1.39$; $t = -1.36$, $p > .01$), thus rejecting H2.

4.2.3. The profile of students' inductive reasoning

The profile of the inductive reasoning ability of the students showed an average score as 0.69 logit (SD = 1.41). In each category, students obtained a mean logit higher than 0 for figure series ($M_{FS} = 0.36$; SD = 1.99), figure analogy ($M_{FA} = 0.69$; SD = 1.66), and number series

($M_{NS} = 1.89$; SD = 1.87), while number analogy has a mean logit slightly lower than 0 ($M_{NA} = -0.01$; SD = 1.76). In addition, the t-test comparison analysis revealed no significant difference between male and female students in inductive reasoning ($M_{male} = 0.78$, $SD_{male} = 1.54$; $M_{female} = 0.65$, $SD_{female} = 1.33$; $t = 1.52$, $p > .01$), thus rejecting H3.

4.2.4. The profile of students' scientific competency

In the overall performance of scientific competency, the students reached a logit score close to 0 logit (M = -0.16; SD = 1.34), meaning that the ability is at an average level. In the SCE category, the students' score is slightly above 0 logit ($M_{SCE} = 0.05$; SD = 1.48), while the SCI category has a lower logit score ($M_{SCI} = -0.43$; SD = 1.64). In terms of gender differences, there is no significant difference between male and female students in the scientific competency test ($M_{male} = -0.21$, $SD_{male} = 1.54$; $M_{female} = -0.13$, $SD_{female} = 1.26$; t = -0.46, $p > 0.05$), thus rejecting H4.

4.2.5. The profile of students' attitude towards science

The overall students' attitude towards science is 0.21 logit (SD = 0.65) slightly above the average 0 logit. The highest score for attitude towards science was found in the SVAL variable (M = 2.36; SD = 2.05) implicating that students tend to give positive responses about the value of science in life and society. For the ENJ variable, the students' response is close to 0 logit (M = 0.49; SD = 2.15) as well as the PAR variable (M = 0.31; SD = 1.39). On the other hand, the ANX variable got the lowest score below logit 0 (M = -0.75; SD = 1.61). In addition, we found no significant difference in attitude towards science between genders ($M_{male} = -0.19$, $SD_{male} = 0.76$; $M_{female} = 0.21$, $SD_{female} = 0.58$; t = -0.66, $p > 0.05$), thus rejecting H5.

4.2.6. The profile of students' science motivation

The students' mean score for SMQ is 1.26 logit (SD = 1.50). Among the five variables in SMQ, GM has the highest average logit (M = 2.43; SD = 2.30), followed by CM (M = 1.76; SD = 2.72), SE (M = 1.70; SD = 2.80), SD (M = 1.64; SD = 2.36), and the last is IM (M = 1.21; SD = 2.03). The students' overall responses in SMQ show a value above average 0 logit, indicating that their science motivation level is approximately high. When comparing science motivation between male and female students, we found a significant difference between genders when females outperformed males ($M_{male} = -0.19$, $SD_{male} = 0.76$; $M_{female} = 0.21$, $SD_{female} = 0.58$; t = -0.66, $p < 0.05$), supporting H6.

4.3. The connection between cognitive and non-cognitive factors in problem solving

To investigate the connection between cognitive and affective factors on problem solving, SEM analysis was performed to confirm the effect of each. The proposed model places domain-general and domain-specific problem solving as dependent variables, while cognitive factors (inductive reasoning and scientific competency) and affective factors (attitude towards science and science motivation) act as independent variables. In addition, gender and socioeconomic status were also added as independent variables for problem solving, cognitive, and affective factors. The model produces an acceptable fit ($\chi 2/df = 6.42$, CFI = 0.92, TLI = 0.90, RMSEA = 0.07, and SRMR = 0.04). We found that science problem solving is associated with MicroDYN ($\beta = 0.21$; $p < 0.01$). In the case of cognitive factors, inductive reasoning directly affects MicroDYN ($\beta = 0.20$; $p < 0.01$) and strongly affects scientific competency ($\beta = 0.75$;

$p < 0.01$), there is no significant direct effect of inductive reasoning on science problem solving (rejecting H7). Furthermore, scientific competency is the only cognitive factor that has a direct effect on science problem solving ($\beta = 0.34$; $p < 0.01$) which supports H8. Scientific competence was also significantly associated with MicroDYN ($\beta = 0.35$; $p < 0.01$). Affective factors, attitudes toward science and science motivation have a direct effect on MicroDYN ($\beta = 0.29$; $p < 0.05$; $\beta = -0.24$; $p < 0.05$) but are not significant for domain-specific problem-solving (rejecting H9 and H10). Regarding gender, we found its significant effects on MicroDYN ($\beta = -0.19$; $p < 0.01$) and science problem solving ($\beta = 0.09$; $p < 0.01$) which supports H11. There is no effect of gender on the cognitive factor, but it has a small effect on attitude towards science ($\beta = 0.09$; $p < 0.01$) and science motivation ($\beta = 0.13$; $p < 0.01$). We found no direct effect of socioeconomic status on problem solving, thus rejecting H12. This variable is associated only with affective factors and inductive reasoning ($\beta = 0.26$; $p < 0.01$).

5. Discussion

The present study explains the assessment results of domain-general and domain-specific problem solving. In addition, the measurement of cognitive factors such as inductive reasoning (IR) and scientific competency (SC), affective factors in attitude towards science (ATSQ) and science motivation (SMQ) was done. The prior analysis of the validity of all measurements confirms that the tests and questionnaires used in this study are valid and reliable for evaluation purposes. Each test and questionnaire show acceptable factor analysis with proficient results in each dimension variable. Rasch analysis also confirms that the items in the tests and questionnaire are in the range of acceptable results (infit MNSQ 0.5–1.5). We find no or negligible risk of item bias in the instrument through DIF analysis concerning gender. The psychometric properties of the instruments show valid evidence and are sufficient to be applied for assessment purposes in the Indonesian context (Wicaksono & Korom, 2023a, 2021).

The assessment results of domain-general and domain-specific problem solving follow descriptive analysis based on the logit parameter. Student performance in both domain-general and domain-specific problem solving was classified as low level with an average logit value below 0 point. In domain-general problem solving, two phases are used for the assessment, knowledge acquisition and knowledge application. Students also show low performance in knowledge acquisition and knowledge application, indicating the difficulty of students in generating knowledge from problem exploration activities with a low chance to apply their knowledge in the new situation. Analysis of gender differences found that males significantly outperformed female students in domain-general problem solving as well as in the knowledge acquisition and knowledge application phase. This finding is consistent with other studies that have reported an advantage for males in domain-general problem solving (Greiff et al., 2018; OECD, 2014; Wicaksono & Korom, 2023c; Wüstenberg et al., 2014). In domain-specific problem solving, the assessment also employs two phases, identifying problem (IP) and generating solution (GS) representing the problem situation in terms of scientific phenomena. The students show low performance in the IP and GS categories of domain-specific problem solving. In contrast to domain-general problem solving, the result of comparison does not indicate a significant difference between domain-specific problem solving of male and female students. Indeed, when it comes to problem solving that requires domain-specific expertise, female students can perform at the same level as male students. This result is also supported by previous studies that found no gender differences in maths problem solving (Gallagher et al., 2000; Guven & Cabakcor, 2013) and science problem solving (Harskamp et al., 2008). Since prior knowledge is required during the completion of domain-specific problem-solving tasks, the similar acquisition of student knowledge during learning and intervention at school can potentially influence their performance in domain-specific problem solving.

In the case of the cognitive factor, the measurement of inductive reasoning reveals that students achieve a score above average, higher than 0 logit. Among the four subtasks of the inductive reasoning test, the students got the highest score in number series and the lowest score in number analogy. The students' performance is more varied when they solve series tasks rather than analogy tasks (Wicaksono & Korom, 2022a), which is confirmed by higher standard deviation. Students have a similar ability to make analogies, but when it comes to making order and finding the pattern of attributes, their ability is diverse. This study shows no gender differences in inductive reasoning tasks. This result is in line with previous studies that found similar performance between male and female students in inductive reasoning (Kambeyo, 2018; Molnár, 2011; Van Vo & Csapó, 2020; Wicaksono & Korom, 2023b). The difference is only found in the number analogy that favours males toward females, but the effect size is relatively small. The other cognitive factor, the measurement of scientific competency, reveals that the students' scores are close to 0 logit on average, presenting a moderate level of scientific competency. In each category of scientific competency, students show better results in explaining scientific phenomena (SCE) than in interpreting and evaluating scientific data or evidence (SCI). Indeed, SCI requires a higher cognitive ability to evaluate and analyse compared to SCI which employs the ability to understand and recall information. In a gender comparison between the ability of male and female students in scientific competency, we found no significant differences. Gender differences in connection with scientific competency are varied as a result of the context and the level of gender equity influences (Cheng et al., 2021). Indeed, the science competency test is built based on the science education curriculum; this non-significant result can be explained by similar cultural contexts and learning experiences throughout the education.

In affective factors, the students' attitude towards science is presented on a moderate level with the highest score in the value of science. This result indicates a higher understanding and acceptance of the influence of science in personal life and communities. The lowest attitude level is found in the anxiety and difficulty variable, indicating that the students perceived fewer negative emotions during science learning. The gender differences analysis does not produce a significant result in the overall attitude towards science of students, but in each category, female students showed a better perception of enjoyment, value of science, and participation in science learning and activities, while male students showed higher anxiety and difficulty level than female students. In the science motivation assessment, students get a high overall score (above 0 logit). Among five categories of science motivation, students show the highest score in grade motivation followed and career motivation, whereas the lowest score is in intrinsic motivation. This result implies that the motivation of the students is mainly driven by external factors such as grades and future careers. Students are motivated to learn science because they want to get high grades rather than their inherent satisfaction and willingness to learn science. However, since the students also showed high intrinsic motivation, self-determination, and self-efficacy, this internal source of motivation also contributes to the learning behaviour of the students.

Regarding the gender comparison, there is a significant difference between the science motivation of male and female students. Female students show higher scores in each category of science motivation than male students, which concludes that they are more motivated to learn science.

Furthermore, the results of this study confirm the effect of cognitive and affective factors, gender, and socioeconomic status on problem solving. In cognitive factors, both inductive reasoning and scientific competency have a significant direct effect on domain-general problem solving, but only scientific competency has a direct effect on domain-specific problem solving. In the case of affective factors, attitude towards science and science motivation both have a direct effect on domain-general problem solving, but no direct effect on domain-specific problem solving. Furthermore, gender has a direct effect on domain-general and domain-specific problem solving, as well as affective factors, while socioeconomic status has a direct effect on inductive reasoning and affective factors. The role of a cognitive factor in problem solving is explained as providing a cognitive basis to understand the problems, apply their knowledge, and make decisions to deal with the problem situation. Furthermore, affective factors will support students' activities and engagement, which potentially improve their strategies in dealing with problem-solving tasks. No significant effect of inductive reasoning and affective factors on domain-specific problem solving raises a concern about their indirect effect and the role of domain-general problem solving and scientific competency as mediators to explain domain-specific problem solving. The result of the study should be explained in light of limitations. In the measurement development of domain-specific problem solving and scientific competency was done in the scope of a topic of natural sciences, which potentially showed different results when used in other disciplines. Instead, the framework for domain-specific problem solving was developed with a focus on two main phases of problem solving, while the other framework encompassed various phases. Similarly, to the scientific competency test, the test framework focused on two main categories that do not fully represent all concepts of scientific competency. The other issue is connected to the inductive reasoning test, which is limited to series and analogy in both figural and numerical form. Other types of inductive reasoning tasks are available for assessment purposes, such as matrix, verbal, and scheme. The further use of this type of test will potentially show more information regarding inductive reasoning measurement. Based on the results of this study, several aspects can be recommended. First, in domain-general and domain-specific problem-solving tests, adding easier tasks is necessary because the majority of items are at a difficult level. Further adaptation of the domain-specific problem-solving task is recommended in other disciplines based on the subjects and context of the studies. The involvement of various domain-specific problem solving (e.g. in science, mathematics, technology, and social studies) will provide a comprehensive profile of the domain-specific problem-solving ability of students, which is important for educational programming and curriculum development. Second, learning instruction and training based on the inductive reasoning process and scientific competency are suggested as an alternative program to foster students' problem-solving ability.

6. Conclusion

The present study provides brief information about the problem-solving assessment in the Indonesian context. The assessment of problem solving in the general and specific domains was carried out with valid evidence of its measurement. The assessment revealed students' problem-solving low-level performance (below 0 logit), while cognitive and affective ability reached a moderate level (slightly above 0 logit). The connection between variables showed that cognitive factors are strongly connected to both domain-general and domain-specific problem solving, whereas affective factors are only connected to domain-general problem solving. In addition, gender is only associated with domain-general problem solving and affective factors, while socioeconomic status is associated with inductive reasoning and affective factors. The results contribute to a better understanding of the factors influencing problem solving and can support the development of curricula and teaching programmes, and the improvement of teaching practice.

References

Aini, R. Q., Rachmatullah, A., & Ha, M. (2019). Indonesian primary school and middle school students' attitudes toward science: focus on gender and academic level. *Journal of Baltic Science Education*, *18*(5), 654–667. https://doi.org/10.33225/jbse/19.18.654

Baker, E. ., & Mayer, R. . (1999). Computer-based assessment of problem solving. *Computers in Human Behavior*, *15*(3–4), 269–282. https://doi.org/10.1016/S0747-5632(99)00023-0

Bat Or, M. (2014). FEATS Problem-Solving Scale in PPAT of children aged 5–6.5 as related to their executive functions and motivation. *The Arts in Psychotherapy*, *41*(1), 27–35. https://doi.org/10.1016/j.aip.2013.11.008

Blum, D., Holling, H., Galibert, M. S., & Forthmann, B. (2016). Task difficulty prediction of figural analogies. *Intelligence*, *56*, 72–81. https://doi.org/10.1016/j.intell.2016.03.001

Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How? *CBE—Life Sciences Education*, *15*(4), rm4. https://doi.org/10.1187/cbe.16-04-0148

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer Science & Business Media.

BSNP. (2013). *Standar isi: standar kompetensi dan kompetensi dasar*. https://kurikulum.kemdikbud.go.id/standar-nasional-pendidikan/

Care, E., Scoular, C., & Griffin, P. (2016). Assessment of Collaborative Problem Solving in Education Environments. *Applied Measurement in Education*, *29*(4), 250–264. https://doi.org/10.1080/08957347.2016.1209204

Cheng, L.-T., Smith, T. J., Hong, Z.-R., & Lin, H. (2021). Gender and STEM background as predictors of college students' competencies in forming research questions and designing experiments in inquiry activities. *International Journal of Science Education*, *43*(17), 2866–2883. https://doi.org/10.1080/09500693.2021.1994167

Chiu, M.-S. (2022). Transcend socioeconomic status constraints to mathematics and science achievement by collaborative problem-solving: The female people-smartness hypothesis. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.944329

Csapó, B. (1997). The development of inductive reasoning: Cross-sectional assessments in an educational context. *International Journal of Behavioral Development*, *20*(4), 609–626. https://www.tandfonline.com/doi/abs/10.1080/016502597385081

Csapó, B., & Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: the eDia system. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01522

Csapó, Benő, & Funke, J. (2017). *The Nature of Problem Solving* (Benő Csapó & J. Funke

(eds.)). OECD. https://doi.org/10.1787/9789264273955-en
Dindar, M., Järvelä, S., Nguyen, A., Haataja, E., & Çini, A. (2022). Detecting shared physiological arousal events in collaborative problem solving. *Contemporary Educational Psychology*, *69*, 102050. https://doi.org/10.1016/j.cedpsych.2022.102050
Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely, M., & Cahalan, C. (2000). Gender Differences in Advanced Mathematical Problem Solving. *Journal of Experimental Child Psychology*, *75*(3), 165–190. https://doi.org/10.1006/jecp.1999.2532
Glaser, R., Chi, M. T. H., & Farr, M. J. (2009). *Measuring complex problem solving: The MicroDYN approach*.
Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159–1176. https://doi.org/10.1002/tea.20442
Gok, T. (2014). Peer instruction in the physics classroom: Effects on gender difference performance, conceptual learning, and problem solving. *Journal of Baltic Science Education*, *13*(6).
Greiff, S., Molnár, G., Martin, R., Zimmermann, J., & Csapó, B. (2018). Students' exploration strategies in computer-simulated complex problem environments: A latent class approach. *Computers & Education*, *126*, 248–263. https://doi.org/10.1016/j.compedu.2018.07.013
Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, *36*, 37–48. https://doi.org/10.1016/j.lindif.2014.08.003
Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic Problem Solving. *Applied Psychological Measurement*, *36*(3), 189–213. https://doi.org/10.1177/0146621612439620
Guven, B., & Cabakcor, B. O. (2013). Factors influencing mathematical problem-solving achievement of seventh grade Turkish students. *Learning and Individual Differences*, *23*, 131–137. https://doi.org/10.1016/j.lindif.2012.10.003
Harskamp, E., Ding, N., & Suhre, C. (2008). Group composition and its effect on female and male problem-solving in science education. *Educational Research*, *50*(4), 307–318. https://doi.org/10.1080/00131880802499688
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
Huang, P.-S., & Chen, H.-C. (2016). Gender Differences in Eye Movements in Solving Text-and-Diagram Science Problems. *International Journal of Science and Mathematics Education*, *14*(S2), 327–346. https://doi.org/10.1007/s10763-015-9644-3
Kambeyo, L. (2018). *Assessing Namibian students' abilities in scientific reasoning, scientific inquiry and inductive reasoning skills*. University of Szeged.
Khine, M. S. (2013). *Application of Structural Equation Modeling in Educational Research and Practice* (Myint Swe Khine (ed.)). SensePublishers. https://doi.org/10.1007/978-94-6209-332-4
Liu, M., Hu, W., Jiannong, S., & Adey, P. (2010). Gender Stereotyping and Affective Attitudes Towards Science in Chinese Secondary School Students. *International Journal of Science Education*, *32*(3), 379–395. https://doi.org/10.1080/09500690802595847
Mayer, R. E., & Wittrock, M. C. (2006). Problem Solving. In Mahwah (Ed.), *Handbook of Educational Psychology* (PA Alexand, pp. 287–303). Erlbaum. https://scholar.google.com/scholar?q=Mayer%2C+R.+E.%2C+%26+Wittrock%2C+M.+C.+%282006%29.+Problem+solving.+In+P.+A.+Alexander+%26+P.+H.+Winne+%28E

ds.%29%2C+Handbook+of+educational+psychology+%28pp.+287-303%29.+Mahwah%2C+NJ%3A+Erlbaum&hl=en&as_sdt=0%2C5&as_
Mihladiz, G., Duran, M., & Dogan, A. (2011). Examining primary school students’ attitudes towards science in terms of gender, class level and income level. *Procedia - Social and Behavioral Sciences*, *15*, 2582–2588. https://doi.org/10.1016/j.sbspro.2011.04.150
Molnár, G. (2011). Playful fostering of 6- to 8-year-old students’ inductive reasoning. *Thinking Skills and Creativity*, *6*(2), 91–99. https://doi.org/10.1016/j.tsc.2011.05.002
Molnár, Gyöngyvér, Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity*, *9*, 35–45. https://doi.org/10.1016/j.tsc.2013.03.002
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. https://nap.nationalacademies.org/catalog/13165/a-framework-for-k-12-science-education-practices-crosscutting-concepts
OECD. (2004). *A Profile of Student Performance in Reading and Science*. OECD Publishing. https://www.oecd-ilibrary.org/docserver/9789264006416-7-en.pdf?expires=1655722645&id=id&accname=guest&checksum=9D08792F6277FEC3D8E25CE40797CD6E
OECD. (2014). *PISA 2012 Results: Creative Problem Solving (Volume V)*. OECD. https://doi.org/10.1787/9789264208070-en
OECD. (2017). *PISA 2015 collaborative problem-solving framework* (pp. 131–188). OECD Publishing. https://doi.org/10.1787/9789264281820-8-en
Oppermann, E., Vinni-Laakso, J., Juuti, K., Loukomies, A., & Salmela-Aro, K. (2021). Elementary school students’ motivational profiles across Finnish language, mathematics and science: Longitudinal trajectories, gender differences and STEM aspirations. *Contemporary Educational Psychology*, *64*, 101927. https://doi.org/10.1016/j.cedpsych.2020.101927
Pásztor, A., Molnár, G., Korom, E. B., Németh, M., & Csapó, B. (2017). Online assessment of inductive reasoning and its predictive power on inquiry skills in science. *17th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI)*, 509.
Reilly, D. (2012). Gender, Culture, and Sex-Typed Cognitive Abilities. *PLoS ONE*, *7*(7), e39904. https://doi.org/10.1371/journal.pone.0039904
Schürmann, L., & Quaiser-Pohl, C. (2022). Out-of-school learning levels prior achievement and gender differences in secondary school students’ motivation. *International Journal of Educational Research Open*, *3*, 100158. https://doi.org/10.1016/j.ijedro.2022.100158
Soto-Ardila, L. M., Caballero-Carrasco, A., & Casas-García, L. M. (2022). Teacher expectations and students’ achievement in solving elementary arithmetic problems. *Heliyon*, *8*(5), e09447. https://doi.org/10.1016/j.heliyon.2022.e09447
Van Vo, D., & Csapó, B. (2020). Development of inductive reasoning in students across school grade levels. *Thinking Skills and Creativity*, *37*, 100699. https://doi.org/10.1016/j.tsc.2020.100699
Wang, M., Wu, B., Kinshuk, Chen, N.-S., & Spector, J. M. (2013). Connecting problem-solving and knowledge-construction processes in a visualization-based learning environment. *Computers & Education*, *68*, 293–306. https://doi.org/10.1016/j.compedu.2013.05.004
Waschl, N., & Burns, N. R. (2020). Sex differences in inductive reasoning: A research synthesis using meta-analytic techniques. *Personality and Individual Differences*, *164*, 109959. https://doi.org/10.1016/j.paid.2020.109959
Wicaksono, A. G. C., & Korom, E. (2022a). The measurement of inductive reasoning skill in

higher education in Indonesia. *JURE 2022: Unpredictable Challenges - Education in a Rapidly Changing World*.
Wicaksono, A. G. C., & Korom, E. (2022b). Review of problem-solving measurement: an assessment developed in the Indonesian context. *Participatory Educational Research*, *9*(1), 116–136. https://doi.org/10.17275/per.22.7.9.1
Wicaksono, A. G. C., & Korom, E. (2023a). Attitudes towards science in higher education: Validation of questionnaire among science teacher candidates and engineering students in Indonesia. *Heliyon*, *9*(9), e20023. https://doi.org/10.1016/j.heliyon.2023.e20023
Wicaksono, A. G. C., & Korom, E. (2023b). Role of inductive reasoning, gender, learning satisfaction, and educational and career preference in predicting scientific competency in high school. *Thinking Skills and Creativity*, *49*, 101376. https://doi.org/10.1016/j.tsc.2023.101376
Wicaksono, A. G. C., & Korom, E. (2023c). The profile of high school students' problem-solving ability in Indonesia. *XIX 19th Conference on Educational Assessment : PÉK 2023*, 26. https://www.edu.u-szeged.hu/pek2023/download/PEK_2023_CEA_2023_absztraktkotet.pdf#page=27
Wicaksono, A. G. C., & Korom, E. (2021). The validation of attitude towards science questionnaire among Indonesian undergraduate students. *EARLI 2021: Education and Citizenship: Learning and Instruction and the Shaping of Futures*, 27.
Wirth, J., & Klieme, E. (2003). Computer-based Assessment of Problem Solving Competence. *Assessment in Education: Principles, Policy & Practice*, *10*(3), 329–345. https://doi.org/10.1080/0969594032000148172
Wüstenberg, S., Greiff, S., Molnár, G., & Funke, J. (2014). Cross-national gender differences in complex problem solving and their determinants. *Learning and Individual Differences*, *29*, 18–29. https://doi.org/10.1016/j.lindif.2013.10.006

RELEVANT PUBLICATIONS

Published papers

Wicaksono, A. G. C., & Korom, E. (2023). Attitudes towards science in higher education: Validation of questionnaire among science teacher candidates and engineering students in Indonesia. *Heliyon*, *9*(9), e20023. https://doi.org/10.1016/j.heliyon.2023.e20023

Wicaksono, A. G. C., & Korom, E. (2023). Role of inductive reasoning, gender, learning satisfaction, and educational and career preference in predicting scientific competency in high school. *Thinking Skills and Creativity*, *49*, 101376. https://doi.org/10.1016/j.tsc.2023.101376

Wicaksono, A. G. C., & Korom, E. (2022). Review of problem-solving measurement: an assessment developed in the Indonesian context. *Participatory Educational Research*, *9*(1), 116–136. https://doi.org/10.17275/per.22.7.9.1

Conferences

Wicaksono, A. G. C., & Korom, E. (2023). The profile of high school students' problem-solving ability in Indonesia. *XIX 19th Conference on Educational Assessment : PÉK 2023*, 26. https://www.edu.u-szeged.hu/pek2023/download/PEK_2023_CEA_2023_absztraktkotet.pdf#page=27

Wicaksono, A. G. C., & Korom, E. (2022). The relationship between science motivation, attitude toward science, and inductive reasoning among science major students in Indonesia. *XVIII 18th Conference on Educational Assessment*, 22.

Wicaksono, A. G. C., & Korom, E. (2022). The measurement of inductive reasoning skill in higher education in Indonesia. *JURE 2022: Unpredictable Challenges - Education in a Rapidly Changing World*.

Wicaksono, A. G. C., & Korom, E. (2021). The validation of attitude towards science questionnaire among Indonesian undergraduate students. *EARLI 2021: Education and Citizenship: Learning and Instruction and the Shaping of Futures*, 27

Paper in progress (to be submitted)

1. Inductive reasoning and scientific competency as mediators for domain-general and domain-specific problem-solving (under review).
2. Measuring science problem-solving in Indonesia by gender and high school grade.