Journal of Contemporary Teacher Education
Vol: 8, 2024, 63-80

Developing and Validating a Test of Higher Order Thinking Skills for Future Teachers

Maimoona Naeem*
Rizwan Akram Rana**

Abstract

Higher-order thinking skills (HOTS) are extremely needed to be possessed by future teachers but there is limited work on the development of suitable measuring instruments to measure HOTS (Zhou et al., 2023). This study is intended to develop an instrument to test the higher-order thinking skills (HOTS) of future teachers at the university level. The instrument named as *Maimoona Higher Order Thinking Skills Test (MHOT)*. The developed test is delimited to two main HOTS: Creative and Critical thinking skills. These HOTS are distinct as well as have inter-dependency. There are a total of fourteen interpretive exercises/items in the instrument. To meet the psychometric criteria the construct validity evidence (n = 150) was gathered through exploratory factor analysis (EFA). The convergent validity evidence was also collected and established by using the College Assessment Academic Proficiency of Critical Thinking (CAAP) test with an average variance of 0.65 and composite reliability of 0.92 between MHOT and CAAP. The inter-rater reliability value of Kappa was 0.724 and the Cronbach's alpha value of the entire instrument was 0.91. The content validation process involved four experts in total, three content experts from the discipline of Education and one from the discipline of Psychology. The scale content validity index (SCVI) was 0.93. Item analysis was also conducted to ensure the discrimination index of the items and that all items were under the acceptable criteria except for two items that were revised.

Keywords: Higher-order thinking Skills (HOTS), Creative Thinking Skills (CRS), Critical Thinking Skills (CTS).

---

* Ph.D. Scholar, Department of Education, University of Management and Technology Lahore.
Email: f2021095012@umt.edu.pk
** Chairperson, Department of Education, University of Management and Technology Lahore Email: rizwan.akram@umt.edu.pk

Introduction

There are now several deliberate efforts to inculcate higher-order thinking skills (HOTS) in the teaching-learning process at every level. There have been multiple justifications to do that because wide-ranged competencies are required for addressing contemporary challenges and we have delineated complex HOTS into some specific and direct observable set of skills (Nagappan, 2002). Higher order thinking generally comprises four primary competencies: critical thinking, creative thinking, problem-solving, reflective thinking and decision-making (Anderson, 2005; Lewis & Smith, 1993).

The evolving social landscape has become a medium of competition among individuals with their experiences, and amassed wisdom (Schmidt, 2022). In response, there is a thriving need for individuals who possess HOTS and a dedication to learning. These individuals integrate both conventional and contemporaneous knowledge, processing information/knowledge with agility, creatively solving problems, collaborating efficaciously in groups, and critically analyzing the problems (Hamlin, 2022; Kroth et al., 2022). Considering such upcoming and occurrent demands, the educational system has to adjust to following 21st century skills or higher-order thinking skills (HOTS) among students by equipping them to grow and compete in a speedily changing society (Tight, 2021a; Tijsma et al., 2020).

Future teachers, as the incoming generation of professionals, play a pivotal role in this undertaking, which necessitates fostering robust HOTS (Naeem & Rana, 2023). It emphasizes the importance of fostering HOTS, especially in terms of critical thinking, problem-solving, creativity, innovation, communication, and collaboration, among future (pre-service) teachers (Brandt et al., 2021; Astuti et al., 2019). These skills are deemed indispensable to cater for the demands of the 21st century post-industrial epoch, also mentioned as Industry 4.0 (Hussain, 2021).

Despite the growing recognition of the importance of HOTS, research focusing on measuring these skills among future teachers remains limited (Abdullah et al., 2016). While elementary, middle, and high school students have been the subject of numerous studies on HOTS measurement, the attention given to future teachers has been relatively scant (Naeem & Rana, 2023; Suparman et al., 2021; Tambunan & Naibaho, 2019). Hence, this study attempts to fill this research void by developing an instrument capable of assessing the HOTS status of future teachers.

Higher-Order Thinking

“Simple knowledge also rests on some historical higher-order thinking. Facts and concepts did not just fall out of the sky—or out of a textbook. They were discovered and debated until they came to be widely held as true, and widely believed” (Brookhart, 2010, p. 06).

HOTS can be traced back to the piece of work of Socrates and Plato. It was considered the cognitive state of expression, logic, arguments, creation, metacognition, judgment, critique, and self-regulation (Resnick, 1987). The concept of "higher-order thinking" more prominently emerged and became widely known through Bloom's taxonomy (Bloom et al., 1956). A theoretical framework formulated for the categorical measurement of learning objectives within the cognitive domain by covering six levels of taxonomy: knowledge, comprehension, application, analysis, synthesis, and evaluation, varying from basic to higher level (Anderson, 2005; Lewis & Smith, 1993). With the supremacy of Bloom’s taxonomy, a question was also urged about what kind of knowledge to practice, that is necessary to complete the tasks and also to measure a precise part of the taxonomy. This holds significant value for teachers when picking out teaching materials. To resolve that in the revised version of taxonomy, the categorizing of knowledge was introduced, and knowledge was dissected into four areas factual, conceptual, procedural and meta-cognitive knowledge (Anderson & Krathwohl, 2001). The focus remains on the content by allowing for a clearer understanding as well as on how to present or adapt it effectively (Miyazaki, 2024; Stayanchi, 2018).

The taxonomy, viewed from a problem-solving standpoint, posits that HOTS are multifaceted by including the skills of generating, inquiring, analyzing and self-regulating, cover multiple criteria, and typically yield a diverse display of problem-solving strategies (Silva Pacheco & Iturra Herrera, 2021). This underscores that HOT is a cognitive function occurring at a higher level (Arievitch, 2020).

It has been reported that the levels of analysis, synthesis, and evaluation are linked to higher-order thinking skills (HOTS) for cognitive development and the rest of the three basic levels are called lower-order thinking skills (Zohar et al., 2003; Stayanchi, 2018; Wiyaka et al., 2020).

HOTS are also potently attached to the constructivist approach that was first assessed through the respective dated sources (Bruner, 1966; Piaget, 1973; Vygotsky, 1978). The approach provides a conducive learning environment to students by involving their mental processes by including creative and critical thinking. The constructivist approach to learning occurs if the learner is involved actively by participating and developing new cognition to be built on pre-existed knowledge or cognition, and the instructor participates as a facilitator (Abosalem, 2016; Anderson & Elloumi, 2004; Nagappan, 2002).

Creative and Critical Thinking as HOTS

Creative and critical thinking are two essential cognitive processes that play critical roles in critique, analysis and innovation. While they share similarities, they also possess distinct characteristics that distinguish them. Sometimes creative and critical thinking skills are viewed as complementary rather than mutually exclusive processes (Ülger, 2016). Both involve cognitive activities with the intent to understand, analyze, and generate ideas. Creative thinking is about inquiring and generating the information or ideas to create novel solutions, investigate possibilities, and think “outside the box”. On the other side, critical thinking relates to analyzing information, evaluating evidence, and reflecting on ideas. It stresses logical reasoning, to identify flaws in arguments (Rosba et al., 2021).

Despite their fluctuation, creative and critical thinking often intersect, with each informing and improving the other. Creative thinking is about the generation of innovative ideas, while critical thinking evaluates and polishes those ideas (Vincent et al., 2019). For instance, in problem-solving contexts, creative thinking prompts people to brainstorm diverse solutions, while critical thinking assists in assessing the practicability and potency of each option. Likewise, in classroom settings, creative thinking inspires future teachers to explore alternate perspectives, while critical thinking supports them to analyze and critically review existing concepts (Karunarathne & Calma, 2024; Wechsler et al., 2018).

The researchers concur with the distinction between generative (creative thinking) and evaluative (critical thinking) elements, with the consensus that both creative and critical thinking typically complement each other and can be categorized under the level of evaluating and creating according to the revised Bloom’s taxonomy (Brookhart, 2010; Stayanchi, 2018).

Methodology

Higher-order thinking skills (HOTS) are premier requirements to be possessed by future teachers, yet research on developing an effective measurement instrument for these skills is limited (Zhou et al., 2023). To adhere to the need, educators have presented the concept of scenario-based exercises as an appropriate assessment to test HOTS. It is an impelling way to achieve the goal of measuring this via the usage of situation-dependent items. Multiple studies mentioned the effectiveness of this type of assessment. The outcomes of numerous studies suggest that scenario-dependent exercises not only improve students' content-related cognition but also improve their general logical abilities. Conversion to scenario-based items can also upgrade the level of complexity and measure higher cognitive skills in an improved way. At the core, this kind of assessment integrates innovative material to measure HOTS (Din & Jabeen, 2014; Salih &Abdelbagi, 2022).

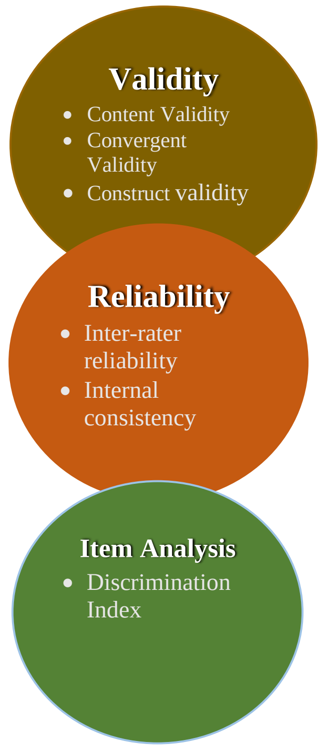
Pilot Test

Piloting is to test an instrument to determine the adequacy of research instruments before its final usage (Creswell & Creswell, 2022). The researcher conducted a small-scale pilot study for three weeks on one hundred future teachers studying in B.Ed. (1.5) program of teacher education in a public sector university of Lahore. The collected data was used for content, construct and convergent validity. Firstly, content validation was conducted. After getting suggestions from experts and revisions of items construct validity was calculated. Then convergent validity was ensured. After the validation procedure, the pilot testing leads to reliability analysis. As far as reliability was concerned two raters were involved in the inter-rater reliability process. After getting a rating from raters an internal consistency analysis procedure was followed and at the end, item analysis was carried out through the item discrimination index.

Psychometric Properties and Analysis

The instrument was pilot-tested and validated through content, construct, and convergent validity. Reliability analysis was conducted using inter-rater reliability and internal consistency measures. Additionally, the item discrimination index was calculated for item analysis.

Figure 1
Summary of Psychometrics Analysis



Analyses and Findings
Validation

Validity is a test measuring what it is supposed to measure (Creswell & Creswell, 2022; Yusoff, 2019). Content validation is a measurement of an intended content domain (Yusoff, 2019). The content validation procedure by Yusoff (2019) was used and Table 1 presents its procedure.

Table 1
Content Validation

Item No.	Expert 1	Expert 2	Expert 3	Expert 4	Experts in Agreement	I-CVI	UA
1	1	0	1	0	2	0.5	0
2	1	1	1	1	4	1	1
3	1	1	1	1	4	1	1
4	1	1	1	1	4	1	1
5	1	1	1	1	4	1	1
6	1	1	1	1	4	1	1
7	1	1	1	1	4	1	1
8	0	1	1	1	3	0.75	0
9	1	1	1	1	4	1	1
10	1	1	1	1	4	1	1
11	1	1	1	1	4	1	1
12	1	1	1	1	4	1	1
13	1	1	1	1	4	1	1
14	1	1	1	1	4	1	1
						0.95 S-CVI Items	0.85 S-CVI UA
Proportion Relevant	0.93	0.93	1.00	0.93			
	S-CVI (Average proportion items by four experts) = 0.95						

"Experts in agreement score" sums up all ratings agreements for each item. "Universal agreement (UA) is score '1' for 100% and 0 for not all the experts agree. Item-content validity index (I-CVI) is the experts in agreement divided by the number of experts. Scale-content validity index of items S-CVI=0.93 is the average score of I-CVI for all items. S-CVI experts are the average proportion relevance scores across all experts. S-CVI UA=0.8 is the average of universal agreement (UA) scores crosswise all items" (Yusoff, 2019, p.53). The criteria provided by Polit et al. (2007) the accepted value for any item is 1.00 in case two to five experts are participating in the content validation process. Thus, all items meet the satisfactory level of content validity except item numbers "1" and "8" (revised later). Experts validated these two items again,

The factor analysis is widely used for construct validation. In this inquiry exploratory factor analysis (EFA) principal axis with varimax rotation was used to calculate the relationship of constructs in 10 items of Maimoona Higher Order Thinking Skills Test (MHOT). The test is based on two main constructs which are creative and critical thinking skills.

Table 2

KMO and Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.80
Bartlett's Test of Sphericity	Approx. Chi-Square	2597.7
	df	91
	Sig.	.000

*($p \leq .05$)

The KMO (Kaiser 1970, 1974) test and Bartlett's Test of Sphericity (Bartlett, 1954) were measured. One fifty future teachers were taken to participate in the test as a sample by following the criterion of Field (2018) that is; a researcher could take any rate of 10-15 subjects for every variable/item. MHOT was comprised of 14 items. The amount of adequacy is .80 which is displaying a strong adequacy of the instrument. The sphericity is also acceptable as .000 is ≤ than .05. All the items loaded under their respective factors with acceptable values and are not less than .05.

Table 3

Factor Loadings

Items	Creative thinking skills	Critical thinking skills
1	.850	
2	.916	
3	.877	
4	.797	
5	.767	
6	.796	
7	.867	
8	.847	
9		.840
10		.811
11		.872
12		.763
13		.667
14		.870

Table 4

Convergent Validity of MHOT1 and CAAP

Factors	*Average Variance*	Composite Reliability
MHOT and CAAP	0.65	0.92
MHOT Creative Thinking and CAAP	0.67	0.91
MHOT Critical Thinking and CAAP	0.63	0.90

The convergent validity was also conducted by calculating the *Average Variance Extracted (AVE)* in use to measure it. Usually, an AVE higher than 0.5 depicts an acceptable value of convergent validity The convergent validity of MHOT was calculated by using the College Assessment Academic Proficiency of Critical Thinking (CAAP CT) Test. CAAP CT is a standardized test and was constructed to measure students' critical thinking skills at the college level. It has 32 items to be solved in forty (40) minutes. Scoring Key is also given by ACT (ACT, 2008).

Reliability

The reliability of an instrument is about the consistency of scores (Creswell & Creswell, 2022). The reliability of the MHOT instrument was assessed using two raters on a sample of thirty -eight future teachers. Due to the raters' convenience and time constraints, thirty -eight tests were chosen for evaluation. This sample size was agreed upon by the raters. Consequently, thirty -eight randomly selected tests were marked by the raters.

After explaining the rubrics and marking criteria to the raters, the raters scored the tests. Cohen’s kappa (for two raters) was used. Cohen’s kappa “k” varies from -1 to +1. (McHugh, 2012).

Table 5

Inter-rater Reliability

Symmetric Measures		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.724	.073	20.981	.000
N of Valid Cases		38			

The two raters marked the same scores for twenty-nine cases of sample and dissimilar scores for nine cases. The agreement means no difference in the scores of both raters and disagreement signifies a difference in scores. The k (Kappa) value of the instrument is 0.724 with 29 agreements and 9 disagreements. It is acceptable as a moderate level of agreement. The percentage of agreement was 76% (29/38) and the percentage of disagreement was 24% (9/38). The value of Cronbach's alpha was 0.915 which is also on the stronger side (McHugh, 2012).

Item Analysis was also carried out by calculating the discrimination index for MHOT. To calculate the discrimination index, the rule of thumb is to organize all the marked tests in descending order and then make two groups based on higher and lower scores. Each group should represent 27% of the total number of students (future teachers).

The total sample was one hundred fifty tests of one hundred fifty future teachers. Thus, the tests were divided into two halves: the upper group consisting of forty students with higher marks, and the lower group consisting of forty students with lower marks. The scores of the lower group were subtracted from the scores of the upper group for every item by following the formula process of Gay (1985. p, 258).

Formula Description

Total Students (T) =150

27% of the Total students (T) of a group = 40

Upper group = 40

Lower group = 40

Formula = The scores of the upper group for each item **-** The scores of the lower group for each item

Table 6 shows the difference between all the items was sufficient enough except for item numbers 3 and 12. Both items showed less difference and were revised later for constructing the final instrument

(MHOT). All the items except two showed enough difference to claim that items are discriminating between higher and lower achievers of the test.

Table 6

Item Analysis with Discrimination Index

Items	Upper Group	Lower Group	Difference	Interpretation
1	118	61	57	Performing effectively
2	114	71	43	Performing effectively
3	128	113	15	Performing ineffectively
4	136	75	61	Performing effectively
5	139	73	66	Performing effectively
6	132	68	64	Performing effectively
7	121	63	58	Performing effectively
8	128	72	56	Performing effectively
9	113	80	33	Performing effectively
10	109	77	32	Performing effectively
11	134	66	68	Performing effectively
12	123	108	15	Performing ineffectively
13	123	71	52	Performing effectively
14	109	64	45	Performing effectively

Summary of the Findings

The societal call for 21st century skills with the use of an effective assessment of future teachers' HOTS as creative and critical thinking skills, forces the development of a valid quantitative MHOT measurement test for future teachers. The validation process was implemented in three ways via calculating content validity, construct validity and convergent. Reliability was calculated in two ways through raters and Cronbach's alpha. The process involved a Scale-content validity index of items as S-CVI=0.95 a scale-content validity index universal agreement (S-CVI UA) is 0.85. Item 1 and 8 were revised by following the suggestions of experts involved in content validation process. EFA item loadings ranged from 0.667 to 0.916. The convergent with CAAP presented average variance as 0.63 and composite reliability as 0.90. Item analysis through the discrimination index showed sufficient difference except for items 3 and 12 which were revised later. The results of reliability showed kappa= 0.72 and Cronbach's alpha = 0.91. The Maimoona Higher Order Thinking

(MHOT) test meets the psychometric criteria to assess the creative and critical thinking skills of future teachers at the university level.

Conclusion

The developed instrument, named the Maimoona Higher Order Thinking (MHOT) test, underwent pilot testing for psychometric analysis. The instrument verified strong validation through content, convergent, and construct validation procedures, affirming its validity. Moreover, reliability measures confirmed the consistency of the MHOT test. Item analysis results specified that the test items effectively discriminated between higher and lower achievers. Therefore, the MHOT is a valid and reliable instrument for measuring the higher-order thinking skills of future teachers at the university level.

References

Abdullah, A. H., Mokhtar, M., Abd Halim, N. D., Ali, D. F., Mohd Tahir, L., & Abdul Kohar, U. H. (2016). Mathematics teachers' level of knowledge and practice on the implementation of higher-order thinking skills (HOTS). *EURASIA Journal of Mathematics, Science and Technology Education, 13*(1). https://doi.org/10.12973/eurasia.2017.00601a

Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education, 4*(1), 1-11.

ACT. (2008*). Collegiate Assessment Academic Proficiency Critical Thinking (CAAP) Technical Handbook.* American College Testing Progra*m.* www.act.org

Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education*. Studies in Educational Evaluation*, *31*(2–3), 102–113.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and* assessing: *A revision of Bloom's taxonomy of educational objectives*: complete edition. Addison Wesley Longman, Inc. https://doi.org/10.1016/j. stueduc.2005.05.004

Arievitch, I. M. (2020). The vision of developmental teaching and learning and Bloom's taxonomy of educational objectives. *Learning, Culture and Social Interaction*. https://doi.org/10.1016/j.lcsi.2019.01.007

Astuti, A. P., Aziz, A., Sumarti, S. S., & Bharati, D. A. L. (2019). Preparing 21st century teachers: Implementation of 4C character's pre-service teacher through teaching practice. *Journal of Physics, 1233*(1).

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals by a Committee of College and University Examiners*. *Handbook I: Cognitive Domain*. New York: Longmans Publishing.

Brandt, J.-. O., Barth, M., Merritt, E., & Hale, A. (2021). A matter of connection: The 4 Cs of learning in pre-service teacher education for sustainability. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2020.123749

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.

Bruner, J. (1966). *Toward a theory of instruction*. Massachusetts. Harvard University Press.

Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches.* (6th ed.). Sage.

Din, A. M., & Jabeen, S. (2014). Scenario-based assessment exercises and the perceived learning of mass communication students. *Asian Association of Open Universities Journal, 9*(1), 93-103.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage.

Gay, L. R. (1985). *Educational Evaluation & Measurement.* Competencies for analysis and application.

Hamlin, M.D. (2022). *Developing self-directed learning skills for lifelong learning* (pp. 209–234). https://doi.org/10.4018/978-1-7998-7661-8.ch012. https://doi.org/10.11591/edulearn.v13i1.11218

Hussain, Z. (2021). Paradigm of technological convergence and digital transformation. The challenges of CH sectors in the global COVID-19 pandemic and commencing resilience-based structure for the post-COVID-19 era. *Digital Applications in Archaeology and Cultural Heritage*, *21*, e00182. https://doi.org/10.1016/j.daach.2021.e00182

Jansen, T., & M¨oller, J. (2022). Teacher judgments in school exams. Influences of students' lower-order-thinking skills on the assessment of students' higher-order-thinking skills. *Teaching and Teacher Education*, *111*, https://doi.org/10.1016/j.tate.2021.103616

Kaiser, H. F. (1970). *A second generation little jiffy*. *Psychometrika*, *35*(4), 401–415.

Kaiser, H. F. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement, 34*(1), 111–117. https://doi.org/10.1177/001316447403400115

Karunarathne, W., & Calma, A. (2024). Assessing creative thinking skills in higher education: deficits and improvements. *Studies in Higher Education, 49*(1), 157-177.

Kroth, M., Carr-Chellman, D. J., & Rogers-Shaw, C. (2022). Formation as an organizing framework for the processes of lifelong learning. *New Horizons in Adult Education and Human Resource Development, 34*(1), 26–36. https://doi.org/10.1002/nha3.20348

Lewis, A., & Smith, D. (1993). Defining higher order thinking. T*heory Into Practice, 32*(3), 131–137. https://doi.org/10.1080/00405849309543588

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276-282.

Miyazaki, R. (2024). *Using Revised Bloom's Taxonomy to Evaluate Higher Order Thinking Skills (HOTS) in Tasks From an 8th Grade English Language Textbook in Japan* (Master's thesis, Southeast Missouri State University.

Nagappan, R. (2002). Using constructivist approach to teach higher order thinking skills. Transforming Teaching Practice to Facilitate Mindful Learning. In *10th International Conference on Thinking* (pp. 15-19).

Naeem, M., & Rana, R. A. (2023). An Investigation about Creative and Critical Thinking Skills of Prospective Teachers: A Preliminary Study. *Bulletin of Education and Research*, *45*(3), 35-47.

Piaget, J. (1973). *To Understand is to Invent.* New York: Grossman. (http://curriculum.calstatela.edu/faculty/psparks/theorists/501const.htm)

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on research methods is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Researchin Nursing & Health*, *30*, 459–467.

Resnick, L. (1987). *Education and Learning to Think*. National Academy Press.

Rosba, E., Zubaidah, S., Mahanal, S., & Sulisetijono, S. (2021). College students' critical thinking skills and creativity. *AIP Conference Proceedings, 2330* (1).

Salih, M., & Abdelbagi, O. (2022). Scenario-Based, Single Best, Multiple-Choice Questions (SB-SB-MCQs) in basic medical sciences. An exploratory study about the staff awareness, knowledge and difficulties encountered. *Journal of Biosciences and Medicines, 10* (9), 79-85.

Schmidt, E. (2022). AI, great power competition & national security. *Daedalus, 151(*2), 288–298. https://doi.org/10.1162/daed_a_01916

Silva Pacheco, C., & Iturra Herrera, C. (2021). A conceptual proposal and operational definitions of the cognitive processes of complex thinking. Thinking Skills and Creativity, 39. https://doi.org/10.1016/j.tsc.2021.100794

Stayanchi, J. (2018). Higher order thinking through Bloom's taxonomy. *Kwansei Gakuin University Humanities Review, 22*, 117-124.

Suparman, S., Juandi, D., & Tamur, M. (2021). Does problem-based learning enhance students' higher order thinking skills in mathematics learning? A systematic review and meta-analysis. 4*th International Conference on Big Data and Education* (pp. 44–51). https://doi.org/10.1145/3451400.3451408

Tambunan, H., & Naibaho, T. (2019). Performance of mathematics teachers to build students' high order thinking skills (HOTS). *Journal of Education and Learning (EduLearn)*, *13*(1), 111–117.

Tibbetts, G. G. (2013). Elements of scientific thinking. *How the Great Scientists Reasoned* (pp. 5–24). Elsevier. https://doi.org/10.1016/B978-0-12-398498-2.00002-3

Tight, M. (2021a). Twenty-first century skills: Meaning, usage and value. *European Journal of Higher Education, 11*(2), 160–174. https://doi.org/10.1080/ 21568235.2020.1835517

Tijsma, G., Hilverda, F., Scheffelaar, A., Alders, S., Schoonmade, L., Blignaut, N., et al. (2020). Becoming productive 21st century citizens. A systematic review uncovering design principles for integrating community service learning into higher education courses. *Educational Research, 62(*4), 390–413. https://doi.org/10.1080/00131881.2020.1836987

Ülger, K., Morsünbül, Ü. (2016). The differences in creative thinking: the comparison of male and female students. *The Online Journal of Counselling and Education*, *5*(4), 1-12. http://www.tojce.com/frontend//articles/pdf/v5i4/1-12-ulger-morsunbulpdf.pdf

Vincent-Lancrin, S. et al. (2019), *Fostering Students' Creativity and Critical Thinking: What it means in school*, OECD Publishing, Paris, https://doi.org/10.1787/62212c37-en.

Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.

Wiyaka, W., Prastikawati, E. F., & Kusumo Adi, A. P. (2020). Higher-order thinking skills (hots)-based formative assessment: a proposed model for language learning assessment. *Vision: Journal for Language and Foreign Language Learning*, *9*(2), 115-130.

Wechsler, S. M., Saiz, C., Rivas, S. F., Vendramini, C. M. M., Almeida, L. S., Mundim, M. C., & Franco, A. (2018). Creative and critical thinking: Independent or overlapping components? T*hinking Skills and Creativity*, *27,* 114-122.

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal, 11*(2), 49-54.

Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity, 48*, 101272.

Zohar, Anat, & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *Journal of the Learning Sciences, 12*(2), 145–181. https://doi.org/10.1207/S15327809JLS1202_1

Citation of this Article: Naeem, M. & Rana, R. A. (2024). Developing and validating a test of higher order thinking skills for future teachers. *Journal of Contemporary Teacher Education, 8*, 63-80.