VALIDITY OF MEASURING CRITICAL THINKING

Bachelor thesis

A research about the validation of the two questionnaires CT-HK and CriTT that are used for measuring the important skills of critical thinking among Applied Psychology students.

02 JUNI 2020

KAREN BOKDAM
Saxion University of Applied Science

Commissioned by:
Saxion University of Applied Sciences,
and Tomsk State University

Thesis written by:
Name: Karen Bokdam
Student: 122164
Email: 122164@student.saxion.nl

Submitted for:
Bachelor of Applied Psychology,
Major Clinical Psychology
at Saxion University of Applied Sciences
Deventer, The Netherlands

Supervised by:
First supervisor: Dr. L.F. Ekkel
Second Supervisor: MSC. A. Winkler
2nd of June, 2020

PREFACE

In this preface page I want to give you the opportunity to get to know something about this paper and something about the writer of this paper. This thesis of validity research is written and performed by Karen Bokdam. I´m an applied Psychology student at the Saxion University Applied Science. In the first three years of my education I learned a lot about diagnostic and applied research in addition to the theoretical courses on clinical and neurological psychology. The assignments for the applied research courses always sparked my interest. Besides the good education about applied research that was given to me by the talented teacher I noticed it was also something I never had major problems with performing. So, for my last college year I did a research about Critical Thinking that already had several studies performed by other students before me. When I heard about the research of critical thinking among psychology students, I got quickly enthusiastic. I already was a big believer that critical thinking is a very important skill to own in different kinds of settings. It's not only a highly recommended skill in the working fields but also very helpful in education and private settings. For example, I have experienced that for problem solving it's helpful to look at the issue from different angles. And when I heard that the test materials that were used by my former fellow students, didn´t had many researches for validation yet, it got my attention right away. So, when I got the opportunity to do this validity research, I took it with both hands.
During this research I was going to a rough time in my personal life. It had a slow start with writing the research plan. And when the real research started, I was diagnosed with ADD. This is an attention disorder that gave me some challenges in my former college years. And one of those challenges was that I could not give my full concentration for a longer time of period. And none of my concentration when I was not interested in the subject. But when someone with ADD finds his or her interest, a 'hyper focus' goes on. You just can't stop what you're doing. I took advantage of this knowledge in this research besides the fact that this paper is written in quarantine time caused by COVID-19. Especially whit the literature research, I got so taken with the theoretical frame, a lot of time was spent on it. Time which luckily given me by this ´smart lockdown´ period. A lot of questions came up in my head. And with these questions and the results of this research, recommendations have been made that can provide a good follow-up study for the validation of measuring critical thinking among psychology students. With this being sad, I can proudly finish my education with this paper.

I would like to express my gratitude to my supervisor Dr. L. Ekkel. He gave me this wonderful opportunity and with patience he took the time to help me with this research. I want to thank him for his feedback, for sparring with me on the subject and results, and specially for his trust in me that I can do a great job on my final assignment to graduate.
I would like to give a special thanks to MSC M. Coopmans for the moral support, the short pep-talks that took place weekly, giving me tips when I got lost in the process and the faith in me to finish my education with this great research.
Also, I would like to give a special thanks to my second supervisor MSC A. Winkler for her enthusiasm being involved with this project and the great feedback she has given me. And finally, my gratitude goes to my friends for giving me the moral support that I needed, my brother that give me feedback and my family for never giving up the faith in me. Without these persons this thesis has not been what it is now.

Karen Bokdam

Haaksbergen, June 2020

ABSTRACT

The purpose of this study is to investigate if the questionnaires CT-HK and CriTT are valid for measuring critical thinking. The skill critical thinking is one of the aspects tutors of Saxion University of Applied Science and Tomsk State University are looking at for the students on the study program Applied Psychology. Over the years the CT-HK and CriTT are used for research about the critical thinking skills owned by students and on which critical thinking skills there can be improvement. There is a lack on the reliability and validity research on the two toolkits. Some studies showed a good reliability of the CT-HK. The CriTT have been used several times for researches but the reliability remained unwritten in many papers.

Experts over the world have agreed on that critical thinking contains 7 dispositions that can be improved for critical thinking skills. The dispositions are 1) inquisitiveness, 2) open-mindedness, 3) systematicity, 4) analyticity, 5) truth-seeking, 6) critical thinking self-confidence and 7) maturity. The CT-HK contains 4 of these dispositions that matches with inquisitiveness, open-mindedness, systematicity and truth-seeking. This makes the CT-HK practically valid for measuring critical thinking. The CriTT only contains 1 disposition that matches with critical thinking self-confidence. In addition, the CriTT also measures the attitudes and beliefs of critical thinking. This makes the questionnaire a suitable supplement for measuring critical thinking. But it is not valid for measuring critical thinking.

For this study an existing database of 62 Applied Psychology students is used to examine the reliability and validity. The results show on overall good reliability for the CT-HK. One of the disposition categories shows an unacceptable reliability. And for the CriTT there also an overall good reliability is indicated and an unacceptable reliability for one category, misconception, that does not match with a critical thinking disposition. The correlation analysis shows a surprising outcome for the CT-HK and CriTT. The correlation is not very high but positive with a no significant difference level. The construct of the two toolkits are completely different but there is a possibility that they correlate.

The validity of the CT-HK and CriTT in this study is partially confirmed when used together for measuring critical thinking. It can be improved with adding another questionnaire or items from another toolkit. Further research of the validity of the questionnaires are necessary to conclude if the questionnaires are valid for measuring critical thinking. In this study the database is to small for the extern validity. In further research more respondents are needed to generalize the results for all Applied Psychology students.

TABLE OF CONTENT

LIST OF TABLES

CHAPTER 1. INTRODUCTION TO CRITICAL THINKING

In the first chapter a brief introduction to critical thinking is made. First, a description of the birth of this research is presented in paragraph 1.1. Followed by a justification of the importance of critical thinking in paragraph 1.2. The main question is written in paragraph 1.3. Sub questions have been made to answer the main questions. Non a less, the objective and background of the objective will be described in the last paragraph 1.4.

1.1 CAUSE

"Critical thinking is the ability to think clearly and rationally about what to do or what to believe." (Lau & Chan, 2004)

Critical thinking (CT) becomes more and more popular these days. It has his expectations to reach out for the top 3 job skills in 2020 (Charlton, 2016). Critical Thinking requires a different kind of skill set. For example, to think clearly and rationally you need to be able to understand the logical connections between ideas, be able to argument and evaluate arguments, be able to solve problems, see the importance of idea's and reflect your own mistakes. But also, be able to reflect on the justification of your own or other beliefs and values. It's not an oncoming thing to have struggles with these skills. Many people find it difficult to handle being criticized or find it easier to just be giving the right answer instead of figuring it out themselves (Lau & Chan, 2016).

With this knowledge's it's understandable why critical thinking is becoming a high requirement in the job industries. And education is not only meant to put the knowledge of one specialty or domain. But also, to improve the skills that are necessary for the future employee. Critical thinking is not just one of them. It's one of the most required skills. Because it can help with acquiring the knowledge, improve the theories and it is crucial for self-reflection (Lau & Chan, 2018).

1.2 CRITICAL THINKING AND EDUCATION

In today´s world critical thinking ability is essential for success (Halx & Reybold, 2005). Critical thinking skills have become an important attribute in college graduates for employers. For a graduate Applied Psychology student various critical thinking skills dimensions into their daily work. This includes analysing complex information in a systematic manner, being open-minded and considering various complex solutions to problems. So, teachers must teach students the development of critical thinking. Jeffrey M. Lederer did a study of critical thinking skills on therapy students with the CCTDI questionnaire. The results are showing that higher education students are more motivated to use critical thinking skills then obtaining the critical thinking abilities. Test materials can be used to examine the effectiveness of educational strategies for improving the disposition for critical thinking among students. By making students aware of their own thinking and how they apply different thinking skills; they are better able to control and improve their thinking (Jones & Ratcliff, 1993, p. 10).

Students of the Saxion University of Applied Science have done several studies among CT within Applied Psychology students of two different countries. They have used the instruments CriTT and CT-HK to measure the critical thinking skills of 64 third year Applied Psychology students. The CT-HK is well established in the international literature while the CriTT is a relatively new questionnaire.

There is not much information about the validation of these two instruments written in the aforementioned appointed studies. For future studies it's good to know if these two instruments are really testing the skills of critical thinking. If we know which instrument, the CriTT or CT-HK, is valid for testing critical thinking, the studies give a better overview of critical thinking skills among the Applied Psychology students.

1.3 RESEARCH QUESTIONS

1.3.1 MAIN QUESTION

In order to carry out this research, a main question have been formulated:

"In order to find out more about critical thinking at Applied Psychology students, which instrument, considering the CriTT or CT-HK, is valid for measuring critical thinking?"

1.3.2 SUB QUESTIONS

To answer the main question, four sub questions have been formulated:
1. What does the literature say about the outcomes of the test results of the CriTT and CT-HK about critical thinking?
2. Which questions of both instruments are valid if you're testing critical thinking?
3. What is the reliability of the CriTT?
4. What is the reliability of the CT-HK?
5. How high is the correlation between the CriTT and CT-HK?

1.4 OBJECTIVE OF THE RESEARCH

With over twenty thousand students, Saxion is one of the largest schools, for higher education, in the Netherlands. They are specialist in applied science and research. It has different departments for different specialties. AMA (School of applied Psychology and Human Resources Management) is one of them. They have educations for applied Psychology and Human Resources. The department AMA is working together with Tomsk State University to improve the critical thinking skills of Psychology students. There already have been different researches in order of AMA. M. Wientjes did a research about critical thinking at second year applied Psychology students at Saxion and Tomsk and used the CriTT and CT-HK. Her research had a low validity and reliability (Wientjes, 2016). R. Vreede did the same research at first year applied Psychology students at Saxion and Tomsk. In this research they say that there is also poor validity and reliability results (Vreede, 2015). Both students hadn't done any research of the validity of the instruments. To improve the critical thinking skills of the students M. Margosyan did her research with COIL at third year students of Applied Psychology. She found out that the reliability of the instruments where good, but the validity was unknown. AMA is wondering which instruments are valid and reliable so the results can say more about critical thinking skills. These results can give a better view on the level of critical thinking students master and how to improve measurement of critical thinking.

CHAPTER 2. THEORETICAL FRAMEWORK

In this chapter a literature research has been written. Different kinds of definitions of critical thinking is presented in paragraph 2.1. Followed by a description of the skills and dispositions of critical thinking in paragraph 2.2. How to measure critical thinking and information about the questionnaires are presented in paragraph 2.3. Some issues with measuring and generalizing critical thinking is discussed in paragraph 2.4. And with the information four hypothesis are made and described in paragraph 2.5.

2.1 DEFINITION

Critical Thinking has different definitions so it can't be defined by just one definition. It's not just one skill that a person can possess. It's about different kind of skills and every individual has one or more skills of critical thinking. Each individual has his own level of critical thinking skills that they master. John Dewey was seen as the 'father' of critical thinking and he called it reflective thinking (Fisher, 2001). He defined reflective thinking as:

'Active, persistent, and careful consideration of a belief or supposed form of knowledge in the light of the grounds which support it and the further conclusions to which it tends.' (Dewey, 1909).

Edward Glaser build on this definition of Dewey and made it critical thinking. With his co-worker Watson he developed one of the most popular critical thinking tests on the world, the *Watson-Glaser Critical Thinking Appraisal.* He defines critical thinking as:

'An attitude of being disposed to consider in a thoughtful way the problems and subjects that come within the range of one's experience; knowledge of the methods of logical enquiry and reasoning; and some skill in applying those methods. Critical thinking calls for a persistent effort to examine any belief or supposed form of knowledge in the light of the evidence that supports it and the further conclusions to which it tends.' (Glaser, 1941).

And then one of the most famous contributors in the research of critical thinking, Robert Ennis. He defines critical thinking in a more widely used definition:

'*Critical thinking is reasonable, reflective thinking that is focused on deciding what to believe or do.'* (Ennis, 1989).

For the final definition, the most used definition of scholar will be used. Richard Paul explained that there are three crucial dimensions of critical thinking; 1) the perfections of thoughts, 2) the elements of thoughts and 3) the domains of thoughts. With these dimensions in mind he's definition of critical thinking is:

'Critical thinking is disciplined, self-directed thinking which exemplifies the perfections of thinking appropriate to a particular mode or domain of thought' (Paul, 1990).

2.2 CRITICAL THINKING SKILLS

Despite different definitions of critical thinking, researchers of critical thinking have agreed on the ability's encompassed by the definition. These agreements include analysing arguments, claims or evidence. Making inferences using inductive or deductive reasoning, judging or evaluating and making decisions or solving problems. This means there are more ability's and skills needed to think critically. Facione (1990) speaks of six kinds of critical thinking skills, namely 1) interpretation, 2) analysis, 3) inference, 4) evaluation, 5) explanation and 6) self-regulation. Paul (1990) had three more skills to add to the critical thinking list: the empirical and conceptional dimension, and assumptions. Experts of the Delpi Report (American Philosophical Association, 1990) describes the ideal critical thinking, attempting these skills, to be habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgements, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results. When an individual owns these skills, it doesn´t consequently implies that he or she is a strong critical thinker (Facione, 2015). You can have those skills but not use them. So an individual has to be more than these skills. Practicing these skills can improve critical thinking attitudes which experts are calling 'dispositions'. Facione (1992) made a list of seven dispositions of critical thinking based on this description. The dispositions contain inquisitiveness, open-mindedness, systematicity, analyticity, truth-seeking, critical thinking self-confidence and maturity. These seven dispositions are agreed by other experts worldwide (Ku, 2010).

The disposition inquisitiveness means one´s intellectual curiosity and desire for learning. The disposition open-mindedness means being tolerant of divergent views and sensitize to the possibility of one's own bias. It means being tolerance and understanding of the beliefs and lifestyles of you own and others. The disposition systematicity means being organized, orderly, focused and diligent in inquiry. The disposition analyticity means prizing the application of reasoning and the use of evidence to resolve problems, anticipating potential conceptual difficulties. It means being alert to the need to intervene. The disposition truth-seeking means being eager to seek the best knowledge in each context, asking questions and being objective about pursuing inquiry. It means doing this without letting own self-interests or one's preconceived opinions weigh in or take it into consideration. With the dispositions critical thinking self-confidence, they mean the trust one places in one's own reasoning processes. And finally, with the disposition maturity it means to be judicious in the person's decision making. A person with the disposition maturity can be described as one who approaches problems, inquiry and decision making with a sense that some problems are necessarily ill-structured, some admit of more options and judgments must be made based on standards and evidence (Facione, 1995).

2.2.1 BLOOM´S TAXONOMY

Worldwide the Bloom´s taxonomy model is used for the development of higher-level thinking skills. Benjamin Bloom collaborates in 1956 with M. Englehart, E. Furst, W. Hill and D. Krathwohl and published a model for categorizing educational goals, ´Taxonomy of Educational Objectives. The original Taxonomy from 1956 consisted of six major categories; Knowledge, Comprehension, Application refers to the use of abstractions and concrete situations, Analysing, Synthesis and Evaluation Engenders. In 2001 a group of cognitive psychologists, curriculum and instructional researchers published a revision of the Bloom´s Taxonomy (Anderson et al., 2001). The title was renamed to 'A taxonomy for

Teaching, Learning and Assessment'. With this title it points more to a dynamic conception of classification. The basic of this taxonomy is the knowledge category. Anderson and Krathwohl (2001) made 'create' the highest category and have renamed the other categories in the following; Knowledge-Remember, Understand-Describe, Explain, Apply, Analyse, and Evaluate. According to Black & Ellis (2010) students should require learning to work at all these categories of thinking but Critical Thinkers would be able to work at the higher categories of the revised taxonomy.

It's not easy to learn to be able to know what kind of category you must appeal to in a situation, without bias and taking views from others, into consideration. And even more difficult to teach. However, there are order categories that can teach students to develop their higher thinking skills and use them to be critical thinkers (Forrester, 2008). Facione (2010) made a list of those skills and a selection is presented in the following list:

- Consideration and evaluation of different points of view
- Open-mindedness
- Development of a logical argument with appropriate evidence
- Identifying the flaws, weaknesses or strengths of an argument and identifying bias in themselves and others
- Establishing priorities or decoding significance
- Analysis of the quality of sources
- Synthesise from a variety of sources
- Deduction – reasoning from the general to the specific
- Induction – reasoning from the specific to the general
- Problem solving, even with previously unknown problems
- Development of criteria for evaluation
- Evaluation of their own decision making and evaluation of their own work and the work of others
- Purposeful, reflective judgement
- Self-regulation.

2.3 MEASURING CRITICAL THINKING

Knowing that the level of the dispositions of critical thinking can give a representation of the thinking skills and critical thinking skills among students, how do you measure those dispositions? There are different kinds of test materials for critical thinking like interviews, programs and questionnaire. Most experts use multiple choice questionnaire for measuring critical thinking. In this research only the latter kind of survey is decribed.

2.3.1 CALIFORNIA CRITICAL THINKING DISPOSITION

The CCTD is a multiple-choice questionnaire developed by Facione, Facione and Sanchez in 1994. The inventory list contains 75-items, forced-choice, 6-point adjective checklist using anchors of 1: Strongly agree and 6: Strongly disagree. The inventory yields an overall score (maximum 450, minimum 75) with the cut point of less than 280 described as "deficient" in critical thinking disposition and greater than 350 described as "outstanding" in critical thinking disposition. (Walsh, 2007). It yields seven subscales, which are truth-seeking, open-mindedness, analyticity, systematicity, confidence, inquisitiveness and maturity. Over the years there were five studies which examined the reliability of the items of the inventory list. Including Kakai (2001) who excluded the problematic items.

Only four scales where left and Kakas renamed them to Intellectual Diligence, Open-mindedness, Nonrelativism and Analyticity. Within these four factors the inventory remained stable. In 2007 Walsh did the last research on the reliability of the items of the CCTD. The study showed a questionable low reliability on the scale's inquisitiveness, open-mindedness and systematicity. Further research of the validity of this inventory list is recommended (Walsh, 2007).

2.3.2 WATSON–GLASER CRITICAL THINKING APPRAISAL TEST

The Watson-Glaser critical thinking appraisal test (is one of the oldest multiple question tests used to measure critical thinking. It's a standard use in a normal version with 40 multiple-choice questionnaire and the short version of 20 multiple-choice items. The items contain the five critical thinking skills 1) Making inferences, 2) Recognizing assumptions, 3) reasoning deductively, 4) interpreting arguments and 5) evaluating arguments. So, this test correctly focuses on some of the critical thinking skills. Unfortunately, the test fails in some skills, like argumentation. The test does not include assessment of the ability to identify informal fallacies (Possin, 2014). Also, only one of the five items have a five-answer option. While the rest of the items have two option answers. This means that the test is sensitive for lucky guesses and is not reliable for retesting.

In 2011 Barnett and Francis did a classroom study for higher level thinking by students of the Educational Psychology course. They used the Watson-Glaser CT Appraisal test in a class of 147 students in the first week of the semester and the last week of the semester. The results showed that the general critical thinking ability significantly increases across the semester. Others study´s like Pascarella and Terenzini in 2005 and Renaud and Murray in 2007, 2008, came with the same results. The explanation for this increase may be the testing effect. There were only four months between the test and retest in the Barnett and Francis study. With these results the statements of Possin are substantiated.

2.3.3 CT-HK

The Hirayama and Kasumi´s Critical Thinking Scale, referred to as CT-HK, is an 18 items questionnaire made in 2004 by Hirayama and Kusumi. The 18 items in this test contains nine items from the California Critical Thinking Disposition Inventory, eight items from the Orientation toward critical thinking scale for Japanese undergraduates and has one new item. This new item focus on assessing four skills of critical thinking. These four skills are 1) inquisitiveness, 2) objectivity, 3) use of a logical approach and 4) reliance on evidence. The four skills match with the four out of seven dispositions that experts agreed on. Inquisitiveness matches with inquisitiveness, objectivity with open-mindedness, use of a logical approach with systematicity and reliance on evidence with truth-seeking. This means that it does not measure the other three dispositions, critical thinking self-confidence, analyticity and maturity.

In 2013 Manalo et al., did a research of a comparison in critical thinking between Asian and Western students. This study showed a good reliability of a α factor 0.71. But validity is unknown. In 2017 Tsuchiya et al. wrote a paper of the study on the characteristics of nurses' eye movements during observation of patients with disturbed consciousness by comparing intuition ability, critical thinking, and clinical experience years. The results show an acceptable reliability of a α factor 0.57 - 0.85.

Also, no reference or indication to a valid testing. In 2018 M. Margosyan researched the improvement of critical thinking skills among applied psychology students by using a collaborative online international learning project. The CT-HK was used for examining the critical thinking skills and the reliability remain to be good with a α factor of 0.73. The validity remains unknown.

In 2019 a teacher Yuya Akatsuka used the CT-HK for examining the coloration of higher thinking skills and the English-speaking skills. Unfortunately, he did not pay attention to the reliability and validity. But the participants did a post- and pre-testing over a period of 5 months. Between these results the mean went up and the standard divination went down. There was no difference between English skill levels (low and high) and the critical thinking skills. Also, the study shows that there is no correlation between English skills and the ability to achieve critical thinking skills. Earlier studies, like the study of Manalo in 2013, shows that there is a coloration between the English proficiency level correlated with achieving critical thinking skills.

2.3.4 CRITT

Stupple et.al. published the critical thinking toolkit (CriTT) in 2017. The toolkit has his purpose to measure the beliefs and attitudes of critical thinking among students. The questionnaire contains 27 Likert scale items that should measure 3 critical thinking skills, which are confidence in critical thinking, valuing critical thinking and misconceptions. These three scales only correspond with one critical thinking disposition, namely critical thinking self-confidence. But it is consistent with the views of Halpern, Stanovich and Bonnefon. There are similarities between the elements of critical thinking in this framework and the dual process theories of thinking and reasoning (Stupple, 2017). Type 1 is the fast, implicit, automatic processes critical thinker and type 2 the analytic processes which are purposeful, self-regulatory, conscious and effortful critical thinker. They agreed that reflective thinking and metacognitive processes are important in these theories of thinking and reasoning.

Stupple included in his study 133 psychology students from the University of Derby to measure the reliability of the questionnaire. Confidence in critical thinking shows a high reliability of a α factor 0.92, valuing critical thinking a good reliability of a α factor 0.79 and on misconceptions a poor reliability of a α factor 0.60. The study showed a significant correlation between the three skills and scores of validities. After a factor analysis and a study showing significant correlations between the 3 skills and scores on Stanovich and West's (1997) Argument Evaluation test, Stupple claims the CriTT to be valid.

The CriTT is used in order studies but there haven't been researched of the toolkit that shows reliability and to be valid. For example, in 2018 Straková and Cimermanová used the CriTT in a study on the development of critical thinking among master students. Aldo, they find 'surprising' results of significant difference between the three skills, which they didn't expect, there have been no attention to the reliability and validity. In 2018 Margosyan used the CriTT next to the CT-HK for the study named earlier above. The study showed an overall good reliability of a α factor 0.88. But the skill Misconceptions showed a poor reliability of a α factor 0.42. For a better reliability the item 10 of skill Misconception has been removed.

2.4 ISSUES WITH GENERALIZING CRITICAL THINKING

Is critical thinking generalisable? It's not only confusing because of the amount of definitions that describes critical thinking. But the expressions 'the generalizability of critical thinking' has two senses (Norris, 1989). An epistemological sense and a psychological sense. The epistemological sense are the principles and standards of critical thinking that are applicable to subjects. The psychological sense is the view of people to apply critical thinking on one subject. In the first chapter it has already been clarified that Lederer claims test can be used to see if students have the critical thinking abilities and that it´s necessary for teachers to teach their students about critical thinking. However, it is not possible to educate critical thinking abilities, it is possible to educate the critical thinking skills or depositions. Someone who thinks critical disponed to seek reasons, try to be well informed, uses credible sources, looks for alternatives, considers seriously points of view other than their own, withhold judgment when the evidence and reasons are insufficient and seeks as much precision as the subject permits, among other activities (Norris & Ennis, 1989).

Robert H. Ennis wrote an article about the dispositions of critical thinking and claims that there are some issues in the conceptualizing of critical thinking dispositions. He describes two issues; a) gender bias and b) subject-specificity issues. With gender bias he means that women are more caring for the worth and dignity of every person. According to Ennis caring for the welfare of others should be a desirable trait for critical thinking. For the second one he comes back at the statement of Glaser (1984); *Many commentators have held that critical thinking is subject-specific.* The critical thinking dispositions are open to the same challenge (Ennis, 1996). But experts agreed on certain dispositions so that means testing the levels of these dispositions can give a certain level of individual critical thinking.

2.5 HYPOTHESES

Based on the previous literature in the Theoretical Framework chapter, the following four hypotheses are formulated.

1) *'The CT-HK is valid for testing critical thinking"*

In the literature research there was no information found saying the CT-HK is valid but the reliability analysis in the study's shows a good reliability. For critical thinking to be tested valid, a questionnaire must contain items that measure the seven dispositions of critical thinking where experts agreed on. CT-HK contains four of these dispositions and the items that measure the dispositions are assembled with items of other questionnaires that showed a good reliability.

2) *'The CriTT is valid for testing critical thinking'*

In studies where they used the CriTT, there has been poor to no research of the reliability of the questionnaire. Although the literature shows that the CriTT only is testing the disposition critical thinking self-confidence, it does test three categories of higher thinking levels. Experts like Facione (2010), Ennis (1996), Paul (1990), etc. agreed that these elements of higher thinking levels are important for critical thinking. The CriTT can be useful for the attitudes of critical thinking among students which can say more about the level of critical thinking of the students when used with other instruments that are testing dispositions of critical thinking. It can give an indication where there can be points of attention by the teachers for learning about critical thinking.

3) *'Items of the CT-HK or CriTT need to be excluded to make the measurement of the construct of critical thinking valid.'*

In previous studies items were deleted for the questionnaires to be reliable. The expectations are that in this research the analysis will give similar results. If the questionnaire has a higher reliability when items are deleted the condition reliability can be achieved for validation.

4) *'Other testing materials/items need to be added for critical thinking to be valid.'*

Experts have agreed on dispositions about critical thinking. The CT-HK contains four of these dispositions and the CriTT only shows the attitudes and beliefs of critical thinking. When other items, or even more questionnaires are included when testing critical thinking it can be valid. Together they are missing two of the seven 'agreed by experts' skills of critical thinking; maturity and open-mindedness.

CHAPTER 3 METHOD

Chapter 3 contains a description of the method that will be used to answer the main and sub questions. A short description is presented in paragraph 3.1. The targets and instruments are explained in paragraph 3.2 and 3.3. In these paragraphs information of the questionnaire that this research handles, are described. This research uses a reliability and correlation analysis. The method and procedure of these analysis can be found in the last paragraph 3.4.

3.1 METHOD

In order to answer the main question, the research will be a congruent validity research. For the sub questions it will be a concept validity research. The research will take place in 3 phases. The first one will be a meaning analysis. Literature research will be applied to answer the sub questions predicatively. Phase two will be the concept validity to answer sub question 2. Data has already been collected by former students of Saxion University of Applied Science. For the data the instruments CriTT and CT-HK have been used in a survey research. For a test to be valid, reliability is a necessary stipulation. So, in phase 3, to answer the sub questions 3 and 4 the data will be analysed with the Cronbach's Alpha formula. For the reliability the test will be analysed apart in total, together and per category. The 4 categories for the CT-HK are Logical Systematic, approach, Inquisitiveness, Objectiveness and Evidence. For the CriTT the 3 categories are Misconceptions, Valuing Critical Thinking and Confidence in Critical Thinking. For answering sub question 5 the formula Spearman's and the formula Pearson's will be used. For this analysis the program IBM SPSS Statistics 24 is used.

3.2 TARGET

The target audience will be the students of the education Applied Psychology at the Saxion University of Applied Science and Tomsk State University. Data has already have been gathered trough an online survey that is offered by email. The data contains the responses of 62 students. The data is originating from first, second- and third-year students. This data will be used again to test the validity of the instruments CriTT and CT-HK. The respondents of 62 students in total will be enough to proceed the analysis that is used for this research. A minimum of 30 respondents is enough for a Spearman formula when the items are ordinal (Baarda & van Dijkum, 2019).

3.3 INSTRUMENTS

The following two instruments/questionnaires will be used for the research: CT-HK and CriTT data.

3.3.1 CT-HK

CT-HK is a scale with 18 items that measures the degree of four skills of critical thinking (Hirayama and Kusumi, 2004). These four components contain logical systematic approach, inquisitiveness, objectiveness and evidence-seeking. The 18 items consist of a five-point Likert scale. One represents strongly disagree and five presents strongly agree. Table 3.1 shows which items measure which skill (Wientjes, 2016). Items 5 and 13 are negative and must be recoded.

Table 3.1 CT-HK item measuring scale

Item scale	Items number
Logical systematic approach	1,2,3,4,5
Inquisitiveness	6,7,8,9,10
Objectiveness	11,12,13,14,15
Evidence-seeking	16,17,18

The CT-HK data will be used to answer the main question and sub questions two, three and four. This instrument has been chosen because it was used for measuring critical thinking among Dutch and Russian students.

3.3.2 CRITT

The CriTT is a scale with 27 items that measures the degree of three skills of critical thinking (Stupple et al., 2017). The three components contain confidence in critical thinking, valuing critical thinking and misconceptions. The 27 items consist of a ten-point Likert scale. One represents 'strongly disagree' and ten represents 'strongly agree'. The table 2.3 shows which items measures which CT skills. No items are negative so no items will be recode.

Table 3.2 CriTT items measuring scales

Item Scale	Items number
Misconceptions	6,10,12,21
Valuing critical thinking	4,5,7,9,16,18
Confidence in critical thinking	1,2,3,8,11,13,14,15,17,19,20,22-27

The CriTT data will be used to answer the main question and the sub questions to, three and four. This instrument has been chosen because it was used to measure the level of critical thinking among Dutch and Russian students.

3.4 DATA ANALYSIS

Before the data is analysed as planned the items need to be recode. After recoding the negative items of both questionnaires, the descriptive statistics can be analysed in SPSS 24. The sub question 3, the reliability of the CriTT and sub question 4, the reliability of the CT-HK, can be analysed with Cronbach's Alpha. For sub question 2, which items are valid of the questionnaires, the same formula will be used. Only this time the categories of the questionnaires will be separated in categories as described before in 3.1 Method. To answer the sub question 5, the correlation between the two questionnaires, the Pearson formula will be used. Because the items are ordinal the Spearman's formula will be used also. For this analysis the total of the two questionnaires will be analysed in his total.

The first hypothesis will be confirmed when the reliability of the CT-HK is reviewed as good or excellent. That means a Cronbach's Alpha of >0.80. The hypothesis will be rejected when Cronbach's Alpha is <0.70 and has a poor reliability outcome. When results of the analysis are in between these goals, the knowledge of the literature will be used to confirm the hypothesis. If the literature says the questionnaire contains enough dispositions of critical thinking the Cronbach's Alpha needs to be acceptable ($0.80 > \geq 0.70$).

The second hypothesis will be approved when the reliability of the CriTT is reviewed as good or excellent. That means a Cronbach's Alpha of >0.80. The hypothesis will be rejected when the Cronbach's Alpha is <0.70 and has a poor reliability. When outcomes of the analysis are in between these goals, the knowledges of the literature will be used to approve the hypothesis. If the literature says the questionnaire contains enough dispositions of critical thinking the Cronbach's Alpha needs to be acceptable ($0.80 > \geq 0.70$).

The third hypothesis will be approved if analysis shows a higher reliability when items are deleted in the questionnaire. That means when the Cronbach's Alpha is <0.70 and will be >0.80 when items are deleted.

To approve the fourth hypotheses and to answer the main question the Spearman's analysis needs to be positive. That means a correlation between .0 and 1 of the two questionnaires. The Spearman´s correlation will measure the strength between the two instruments. The strength of association can be in a single value between -1 and +1. A correlation of -1 is a very negative correlation and +1 a high positive correlation. This correlation must be significant ($p = <0.01$) to approve the hypotheses. That will mean that the probability of the correlation is within the 99% chance of being true (Cohen, 1988). When the correlation is negative or not significant, the hypotheses will be rejected according to COTAN (2010). The literature and outcomes of the approved hypotheses will be used to answer the main and sub questions.

CHAPTER 4. RESULTS

In chapter 4 the results of the IBM SPSS Statistics analysis will be presented in paragraph 4.1 with tables for the descriptive statistics. The Cronbach´s Alpha, Spearman, Pearson's of the questionnaires CT-HK and CriTT in total, and each category will be described and presented in tables. They can be found in paragraph 4.2 and 4.3. In this chapter Critical Thinking is described as the abbreviation CT in the tables.

4.1 DESCRIPTIVE STATISTICS

The procedure of the research is already described in chapter 3. The database used for this analysis are the 62 Applied Psychology students' respondents which were 11 male and 51 female students. The gender statistics are shown in table 4.1.

Table 4.1 Gender statistics CT-HK and CriTT

	Male	Female	Total
Students	11	51	62
Valid percent %	17.7	82.3	100

The mean given answer in the CT-HK was 'agree', number 4 on the five-point Likert scale. The results show a low deviation on the mean. In the CriTT the mean answer was 6,5 but had a higher deviation of this mean. It's showing a high-level difference between answers given by the students. In table 4.2 the statistics of the CT-HK are shown and in table 4.3 the statistics of the CriTT can be read. For the categories there are no striking results compared with the total statistics of the questionnaire. Misconceptions have a lower mean of 5.7 compared with the 6.5 of the CriTT and the deviation is higher with a 1.34. The exact numbers of the statistics per category are shown in table 4.4 for the CT-HK and for the CriTT in table 4.5.

Table 4.2 Statistics of the entire CT-HK questionnaire

	N	Mean	Median	Std. Deviation
Students	62	3.9486	4.0000	0.39513

Table 4.3 Statistics of the entire CriTT questionnaire

	N	Mean	Median	Std. Deviation
Students	62	6.4655	6.5556	0.96161

Table 4.4 Statistics per Category of the CT-HK

	N	Mean	Median	Std. Deviation
Logical Systematic	62	3.9250	4.0000	0.53114
Inquisitiveness	62	4.2000	4.3000	0.65186
Objectiveness	62	3.8113	3.8000	0.50728
Evidence seeking	62	3.7742	3.6667	0.71102

Table 4.5 Statistics per Category of the CriTT

	N of students	Mean	Median	Std. Deviation
Misconceptions	62	5.7056	5.7500	1.39196
Valuing CT	62	6.9194	7.0833	1.27743
Confidence in CT	62	6.4845	6.6066	1.11196

4.2 RELIABILITY ANALYSIS

The results of the reliability analysis show a good Cronbach's Alpha for the entire CT-HK (α=0.813). This result is given when no items were deleted. The items for Logical systematic give result of a poor Cronbach's Alpha (α=0.55). When item 5 is deleted the category shows questionable reliability. The category Objectiveness shows an unacceptable reliability of α=0.462. When item 15 is deleted in this category the reliability will become questionable (α=0.610). The reliability analysis of the CT-HK and the four categories are shown in table 4.6.

Table 4.6 Reliability analysis of the CT-HK and per Category of the CT-HK

	N of items	Cronbach's Alpha	Internal Consistency
Logical Systematic	5	0.577	Poor
Inquisitiveness	5	0.801	Good
Objectiveness	5	0.462	Unacceptable
Evidence Seeking	3	0.618	Questionable
Total CT-HK	18	0.813	Good

Note. Internal consistency according to COTAN (2010)

The results of the reliability analysis of the CriTT can be found in table 4.7. Cronbach's Alpha for the entire CriTT is good (α=0.896) with no items deleted. When item 10, 6 and 4 are deleted in the analysis, the results show an excellent reliability of α=0.912. The category Misconceptions shows an unacceptable reliability of α=0.487. When items are deleted from the category Misconceptions the Cronbach's Alpha only goes down and becomes more unacceptable. If the 4 items of Misconceptions are deleted the reliability of the entire CriTT shows a Cronbach's Alpha of α=0.909 which means it will have an excellent reliability review.

Table 4.7 Reliability analysis of the CriTT and per Category of the CriTT

	N of items	Cronbach's Alpha	Internal Consistency
Misconceptions	4	0.487	Unacceptable
Valuing CT	6	0.671	Questionable
Confidence in CT	17	0.920	Excellent
Total CriTT	27	0.896	Good

Note. Internal consistency according to COTAN (2010)

Because the questionnaires are both used together now for the research of critical thinking among Applied Psychology students, the reliability analysis was also used if the questionnaires are put together. The results of a poor reliability can be found in table 4.8.

Table 4.8 Reliability analysis when both questionnaires are put together

	N	Cronbach's Alpha	Internal Consistency
CT-HK + CriTT	45	0.504	Poor

Note. Internal consistency according to COTAN (2010)

4.3 CORRELATION ANALYSIS

The Pearson and Spearman formula are used for the correlation analysis. The results showed a significant difference at the 0.01level and a positive correlation between the CT-HK and CriTT. In table 4.9 the exact numbers of these correlations can be found.

Table 4.9 Correlation analysis of the CT-HK and CriTT

	N of students	Pearson	Spearman	Sig. 0.01 level
CriTT and CT-HK	62	0.479	0.451	0.000

CHAPTER 5. CONCLUSION, DISCUSSION AND RECOMMENDATIONS

In this chapter the main and sub questions will be answered in the first paragraph, conclusion. Chapter 4, the results and chapter 2, theoretical frame, will be used to answer those questions. Some discussion points are described in paragraph 5.2. With these answers' and discussion points, recommendations can be found in the third and last paragraph of this chapter.

5.1 CONCLUSION

5.1.1 SUB QUESTION 1

What does the literature say about the outcomes of the test results of the CriTT and CT-HK about critical thinking?

The meaning of this question is to find out if the literature agrees with the saying that the questionnaires give results that say something about critical thinking. In the literature research seven dispositions of critical thinking are found that the experts of critical thinking all agreed on. The CT-HK measures four of these dispositions' inquisitiveness, systematicity, truth-seeking and open/mindedness. The CriTT contains only one, critical thinking self-confidence. Therefore, together the questionnaires contain a big amount of five out of seven dispositions of critical thinking. Besides the one disposition CT truth-seeking the CriTT also measures the attitudes and beliefs of the respondents. This is in line with what experts claim that higher thinking levels will show more and better skills to think critically. So, the questionnaire does not only measure one disposition of critical thinking but also the higher thinking level skills for critical thinking.

For the reliability and validity of the questionnaire the literature shows a neglect in the studies. The CT-HK have too little research to say if the results are reliable. Reliability is required for a test to be valid. The literature doesn't say enough about the reliability of the CT-HK and therefore also doesn't say a lot about the validity. And with validity is meant if the CT-HK really is testing critical thinking. The reliability and validity of the CriTT also shows a neglect in the literature. The questionnaire is used many times for different studies, but the studies say nothing about the reliability or validity. The maker of the toolkit, Stupple (2017), claims that the toolkit is valid after a study with 133 psychology students.

5.1.2 SUB QUESTION 2

Which questions of both instruments are valid if you're testing critical thinking?

Comparing with the literature the CT-HK and the category confidence in critical thinking from the CriTT contains the right questions for testing the five dispositions of critical thinking that these tests are measuring. In the data analysis some questions are not reliable enough for testing and so not valid for testing critical thinking. The reliability analysis shows that if item 15, 'It concerns me that I might have biases that I am not aware of' is deleted the reliability of the category Objectiveness becomes substantially higher. This question can be interpreted as an attitude about high level thinking and not as a skill of critical thinking. If there are biases it might say something about the ability of critical thinking but then the question is asked in the wrong way and must contain more information. When testing critical thinking this question is not suitable for a 5 point-Likert questionnaire.

The entire category Misconception can be deleted out of the questionnaire CriTT for testing critical thinking. It has no correspondence with the literature agreed on dispositions of critical thinking and shows an unacceptable Cronbach's Alpha in the analysis. Also the items 4, 'Critically thinking is particularly important in psychology', 6, 'When there is a very strong relationship between two variables we can claim that one causes the other' and 10, 'I prefer to do things where there is a quick answer' are negatively important for the reliability of the CriTT. When these items are also deleted, the test has an excellent reliability. Especially number 6 and 10 are items that say more about the biases for critical thinking than critical thinking skills.

Including the literature research with the current research outcomes hypothesis 3 'Items of the CT-HK or CriTT needs to be excluded to make the measurement of the construct of critical thinking valid' is confirmed. Even more for the CriTT than the CT-HK.

5.1.3 SUB QUESTION 3

What is the reliability of the CriTT?

This question is already a little bit answered with sub question 2. The data analysis shows a good reliability of the entire CriTT. In addition, when the items 4, 6 and 10 are deleted, the reliability is even excellent. And that's what we want when we are measuring. The category Misconceptions shows the lowest reliability of the CriTT. So, it's no surprise that the entire reliability of the CriTT increases when the 4 items of this category are deleted. With these results and the information in the literature research hypothesis 2 'The CriTT is valid for testing critical thinking' is partially in accordance with the prediction from the literature research.

5.1.4 SUB QUESTION 4

What is the reliability of the CT-HK?

The reliability analysis for the CT-HK shows a good reliability of the entire CT-HK. And with the literature research that says that the CT-HK contains four of the seven dispositions of critical thinking, hypothesis 1 'The CT-HK is valid for testing critical thinking' is partially confirmed. With this information and the answer of sub question 3, the hypothesis 4 'Other testing materials/items need to be added for critical thinking testing to be valid' is confirmed. If 3 more dispositions are added to the questionnaires, it will contain the entire 7 dispositions of critical thinking. This has been tried out with the CriTT in previous studies and in this research.

5.1.5 SUB QUESTION 5

How high is the correlation between the CriTT and CT-HK?

The correlation analysis shows a positive correlation between the CT-HK and the CriTT. It has a significant difference at 0.01 level. That means there is an excellent chance in 99% that the difference in outcomes would not be observed if the intervention had no benefit whatsoever (Cohen, 1988). So, if the zero hypothesis is true, the probability of the results is no more than 1%. The zero hypothesis is the thesis that there is no correlation between two variables and there is no statistical coincidence. And with the results of this research's results the zero hypothesis is rejected with an 0.01 level significant. The correlation results for the Pearson and Spearman both a 'half' perfect correlation so can be called moderate. This is a surprising outcome because the expectations were that there is no correlation between the CT-HK and the CriTT. They are both used for measuring critical thinking but they both are measuring different dispositions of critical thinking. That the correlation is moderate doesn't mean that there is really a causal context (Cohen, 1988). It means that there is a possibility

that the CT-HK is coherent with the CriTT. Or in other words: the dispositions inquisitiveness, systematicity, truth-seeking and open/mindedness coherent with critical thinking self-confidence and the beliefs and attitudes of critical thinking.

5.1.6 MAIN RESEARCH QUESTION

"In order to find out more about critical thinking at Applied Psychology students, which instrument, considering the CriTT or CT-HK, is valid for measuring critical thinking?"

The hypothesis 1 'The CT-HK is valid for testing CT' and hypothesis 2 'The CriTT is valid for testing critical thinking' are partially confirmed. The results of the analysis show a good or excellent reliability. They are valid for testing the dispositions that matches the literature. They are not valid for the overall dispositions that the experts agreed on for critical thinking. The hypothesis 3 'Items of the CT-HK or CriTT need to be excluded to make the measurement of the construct of critical thinking valid' and hypothesis 4 'Other testing materials/items need to be added for critical thinking to be valid' are confirmed. The analysis results show that the reliability will grow higher when items are deleted. And to make measuring critical thinking valid, all 7 dispositions needs to be included in test materials. This means that 2 dispositions need to be added from other materials/items to make measuring critical thinking valid with the CT-HK and CriTT.

With the answers to the sub questions and the hypotheses being confirmed or partially confirmed the main question can be answered. In chapter 2, in the literature research, it's explained that critical thinking has seven dispositions and the CT-HK contains four of them. The entire questionnaire has a good reliability. So, the CT-HK says more about critical thinking then the CriTT. The CriTT contains one disposition out of seven. It contains three categories, among which one category is not reliable. But when you delete the unreliable category Misconceptions, the entire reliability will be excellent. There is a positive correlation between the two questionnaires. Therefore, the two questionnaires are saying more about critical thinking when taken together for testing critical thinking among Applied Psychology students.

5.2 DISCUSSION

The results of this research showed that the questionnaire CT-HK and CriTT are partially valid for testing critical thinking among Applied Psychology students when used together. There still two out of seven dispositions that are missing which experts agreed on for critical thinking. The two missing dispositions are maturity and analyticity. Abilities are skills you have or you don't. So critical thinking abilities cannot be learned and taught by teachers. However, the dispositions are skills that can be learned, so to be taught by teachers. It can be claimed that the two missing dispositions, maturity and analyticity, need to be included in testing critical thinking.

The CriTT shows excellent reliability when the four items of Misconceptions are deleted. This means only fourteen questions remained in this questionnaire. For a test instrument to be homogeneity more items are desirable (Drenth & Sijtsma, 2006). The test is made as a toolkit for measuring beliefs and attitudes of critical thinking. The test does not say enough about the critical thinking skills on his own. But multiple experts agreed that reflective thinking and metacognitive processes are important in theories of high thinking levels and reasoning. The CriTT may not be acceptable for measuring critical thinking but it is a great addition to know more about learning critical thinking skills.

Robert H. Ennis (1996) claims that caring for the welfare of others is a desirable trait for critical thinking. One of the biases in measuring critical thinking is gender. Women have this desirable trait more than men have. In the current research most respondents were female. This might be of influence on the results of measuring critical thinking skills. If the gender statistics would be reversed, is the outcome still the same as it is now?

For this research a data of 62 respondents is used. The data for the CT-HK and CriTT is collected with an online survey that was send through email to 230 third year psychology students in total at that time (Margosyan, 2018). For the generalizability of the conclusions form this research over all the third-year psychology students according to the theories of Cohen, the respondents must be higher if the same significant difference level are used that this research has used. And let's say that every year, the first, second, third and fourth year of the Applied Psychology education is the same at a total of 230 students per year. The overall total will be 920 students. And then we need a sample size that's even larger (Cohen, 1988). It's important to understand that this is an example. The total Applied Psychology students is variable every year. Considering students can stop early because of a bad study advice or choose another education and not every year the same amount of people register to start this education. For the validity of the two questionnaires CT-HK and the CriTT 62 respondents will be enough. But for the extern validation or generalizability of the two questionnaires there need to be a larger sample size. However, the sample size does represent and matches with the description of the target group (Verhoeven, 2007).

5.3 RECOMMENDATIONS

The CT-HK can be used as a valid questionnaire for critical thinking but does not measure all the dispositions of critical thinking. It's still missing three dispositions. There for another questionnaire is recommended for measuring critical thinking. The two missing dispositions are maturity and analyticity if you take the questionnaire CriTT for the third missing disposition, critical thinking self-confidence. Test instruments which are measuring these two dispositions can be useful for measuring critical thinking skills. The California Critical Thinking Disposition has the two missing dispositions skills categories. The literature research shows that these two categories are reliable. The study of Kakai (2001) already removed the problematic items and made a new version of the questionnaire with four categories left which contains the two missing dispositions for the CT-HK. This version can be used for further research about including questionnaires or items in the CT-HK and the CriTT. A second option can be to make a new test by putting the four dispositions of the CT-HK, the disposition of the CriTT and the two additional dispositions of the CCTD together and perform a factor-analyse. The factor analysis is an often-used method for construct validity and a good option for analysing the meaning of a test. With this method a multitude of information can be summarized in a smaller number of dimensions, with so as little as possible information getting lost.

The CriTT does not measure critical thinking and has a small number of items to be valid but is a useful addition for measuring the construct of critical thinking. It's a valid questionnaire for testing beliefs and attitudes of critical thinking abilities when items 4, 6 and 10 are deleted. The CriTT is not only a good addition to the CT-HK but can also be added on other questionnaires when testing critical thinking within education. For the reliability of the CriTT and of the CT-HK can be further examined by deleting the items that gave a higher reliability in this research and redo the correlation analysis.

For the CriTT items 4, 6 and 10 can be deleted or the 4 items that represent the category Misconception. For the CT-HK item 15 can be deleted. With the deleted items a new comparison can be made with the Spearman formula to found out if the reliability can be improved with these actions.

A test retesting research is recommended for the validation of the CT-HK. The CriTT even needs it for the reliability. Not many previous studies have payed attention to the reliability of the CriTT. This study shows a good reliability, but implementation factors can have influence the results of this research respondents. Test retesting can be a good way to say more about the reliability and validation results. For the best results it is recommend considering the time between the two testing moments. Dispositions can be learned, and the test can give the respondent an incentive to think. An option for this problem can be two different samples. For the validity of the questionnaires a sample of 30 respondents will be enough. But if the generalization for the entire Applied Psychology students are being counted in, the recommendation is to maintain around 25% of the total students of the entire study program.

REFERENCES

Amer, A. (2006), Reflections on Bloom's Revised Taxonomy, *Electronic Journal of Research in Educational Psychology,* 4 (1), 213-230. Retrieved on 7 April 2020, from: https://www.redalyc.org/articulo.oa?id=293123488010

Armstrong, P. (2016), *Bloom's Taxonomy (teaching guide).* Vanderbilt University Center for Teaching, Nashville.

Akatsuka, Y. (2019). Awareness of critical thinking attitudes and English language skills: The effects of questions involving higher-order thinking. *Journal of Pan-Pacific Association of Applied Linguistics,* 23 (2), 59-84. Doi: https:// doi.org/10.25256/PAAL.23.2.4

Baarda, B. & Dijkum, van C. (2019), *Basisboek Statistiek met SPSS: Handleiding voor het verwerken en analyseren van en rapporteren over (onderzoeks) gegevens* (6th ed.).
Groningen/Utrecht, The Netherlands: Noordhoff Uitgevers bv.

Barnett, E. (2011) & L. Francis (2011), Using higher order thinking questions to foster critical thinking: a classroom study, *Educational Psychology, An International Journal of Experimental Educational Psychology,* 32 (2), 201-211.
Doi: https://doi.org/10.1080/01443410.2011.638619

Browne, M. N. & Keeley, S. M. (2005). *De juiste vragen stellen. Praktische handleiding voor kritisch denken*. Eindhoven, The Netherlands: Uitgeverij Rob van der Peet.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences.* Lawrence Erlbaum Associates, New York USA. Retrieved on 22 April 2020, from: http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf

Ennis, R.A. (1996). Critical Thinking Dispositions: Their Nature and Assessability. *Informel Logic,* 18 (2), 165-182. Doi: https://doi.org/10.22329/il.v18i2.2378

Ever, Lucassen, Meijer & Sijtsma (2010). *COTAN Beoordelingssysteem voor de kwaliteit van tests.* Retrieved on 6 May 2020, from: https://www.psynip.nl/wp-content/uploads/2016/07/COTAN-Beoordelingssysteem-2010.pdf

Drenth, P.J.D. & Sijtsma, K. (2006). *Testtheorie: inleiding in de theorie van de psychologische test en zijn toepassingen* (4th rev. ed.). Houten, The Netherlands: Bohn Stafleu van Loghum.

Facione, P.A. (2015). *Critical Thinking: What it is and why it counts.* Hermosa Beach, CA: Measured Reason LLC. Retrieved on 22 January 2020, from: https://www.researchgate.net/publication/251303244_Critical_Thinking_What_It_Is_and_Why_It_Counts

Facione, P.A., Sánches, C.A., Facione, N.C. & Gainen, J. (1995). The disposition toward critical thinking. *Journal of General Education,* 44 (1), 1-25. Retrieved on 27 February 2020, from: https://www.jstor.org/stable/27797240

Fischer, A. (2001). *Critical Thinking, an introduction* (2nd ed.). Retrieved on 11 march 2020, from: http://assets.cambridge.org/052100/9847/sample/0521009847ws.pdf

Glaser, E. M. (1941). *An experiment in the Development of Critical Thinking*. Teacher's College. New York: Columbia University. Retrieved on 26 February 2020, from: https://www.nactateachers.org/images/Dec10_1_What_Is_Critical_Thinking.pdf

Harms, A. (2016). *Critical Thinking. Learning individually versus collaboratively: a research in Dutch and Russian students* (thesis). Deventer: Saxion University of Applied Science.

Hirayama, R., & Kusumi, T. (2004). Effect of critical thinking disposition on interpretation of controversial issues: Evaluating evidence and drawing conclusions. *The Japanese Journal of Educational Psychology,* 52 (2), 186-198. Doi: https://doi.org/10.5926/jjep1953.52.2_186

Howit, D. & Cramer, D. (2007). *Introduction to Statistics in Psychology* (3rd ed.). Amsterdam, The Netherlands: Pearson Education Benelux.

Kakai, H. (2003). Re-examining the factor structure of the California Critical Thinking Disposition Inventory. *Perceptual and Motor Skills,* 96, 435-438.
Doi: https://doi.org/10.2466/pms.2003.96.2.435

Kelly, Y.L. Ku. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity,* 4, 70–76.
Doi: https://doi.org/10.1016/j.tsc.2009.02.001

Kelly, Y.L. Ku. & Irene, T. Ho. (2010). Dispositional factors predicting Chinese students critical thinking performers. *Personality and Individual Differences, Elsevier,* 48, 54-58.
Doi: https://doi.org/10.1016/j.paid.2009.08.015

Lai, E.R. (2011). *Critical Thinking: a literature review* (research report). Pearson Research Reports. Retrieved on 4 March 2020, from: https://pdfs.semanticscholar.org

Lau, J., & Chan, J. (2018). *What is critical thinking?* Retrieved on 11 January 2020, from: http://philosophy.hku.hk/think/critical/ct.php

Lederer, J.M. (2007). Disposition towards critical thinking among occupational therapy students. *American Journal of Occupational Therapy,* 61, 519-526.
Doi: https://doi.org/10.5014/ajot.61.5.519

Manalo, E., Kusumi, T., Koyasu, M., Michito, Y. & Tanaka, Y. (2013). To what extent do culture-related factors influence university students' critical thinking use? *Thinking Skills and Creaticity,* 10, 121-132.
Doi: https://doi.org/10.1016/j.tsc.2013.08.003

Paul, R. (1990). *Critical Thinking: What Every Person Needs to Survive in a Rapidly Changing World.* Rohnert Park, CA: Center for Critical Thinking and Moral Critique Sonoma State University. Retrieved on 11 March 2020, from: https://jgregorymcverry.com

Paul, R. & Elder, L. (2008). *The miniature guide to Critical Thinking*. Tomales, CA: The Critical Thinking Foundation.

Renaud, D. & Murray G. (2008). A comparison of a subject-specific and a general measure of critical thinking. *Thinking skills and creativity,* 3 (8), 86-93.
Doi: https://doi.org/10.1016/j.tsc.2008.03.005

Stupple, E.J.N. (2017). Development of the Critical Thinking Toolkit (CriTT): A measure of student attitudes and beliefs about critical thinking. *Thinking Skills and Creativity,* 23, 91-100. Doi: https://doi.org/10.1016/j.tsc.2016.11.007

Thomas, T. (2011). Developing First Year's Students Critical Thinking Skills. *Asian Social Sience,* 1 (4), 26-35. Doi: doi:10.5539/ass.v7n4p26

Verhoeven, N. (2007). *Wat is onderzoek? Praktijkboek methoden en technieken voor het hoger onderwijs* (2nd rev. ed.). Amsterdam, The Netherlands: Boom Onderwijs.

Vreede, R. (2015). *Critical Thinking: a comparison between Russia and the Netherlands* (thesis). Deventer, The Netherlands: Saxion University of Applied Science.

Wientjes, M. (2016). *Implementation of Critical Thinking in education, in Russia and The Netherlands* (thesis). Deventer, The Netherlands: Saxion University of Applied Science.

Walsh, C.M. (2007). California Critical Thinking Disposition Inventory: Further Factor Analytic Examination. *Sage Journal, Perceptual and Motor Skills,* 104 (1), 141-151.
Doi: https://doi.org/10.2466/pms.104.1.141-151

Walsh, C.M. & Seldomridge, L.A. (2006). Measuring Critical Thinking One Step Forward, One Step Back. *Nurse Educater,* 31 (4), 159-162. Retrieved on 11 March 2020, from: http://www.cncimperia.it/doc/Measuring_criticalthink.pdf

Tsuchiya, R., Shimizu, M., Mikami, K., Aizu, K., Hirakawa, Y., Takase, S. and Nishizawa, Y. (2017) The Characteristic of Nurses' Eye Movements during Observation of Patients with Disturbed Consciousness. *Open Journal of Nursing* , 7, 1502-1514.
Doi: https://doi.org/10.4236/ojn.2017.712105

APPENDIX A.

THE CRITICAL THINKING SCALE OR CT-HK.

Response Options
1: Disagree
2: Mildly Disagree
3: Undecided
4: Mildly Agree
5: Agree

1. I try to provide logical explanations so that everyone can understand and agree with what I mean.
2. I try to develop orderly plans to address complex problems.
3. I try to clarify the assumptions and definition of terms in arguments.
4. I try to organize and clarify the thoughts that others have expressed by using my own words.
5. When I have to deal with something really complex, I tend to panic.
6. I want to meet different kinds of people, and to learn a lot from them.
7. I think that it is important to learn about the thinking styles of people from other countries.
8. I am interested in people with different ideas from me.
9. I want to study about other cultures.
10. Studying new things all my life would be wonderful.
11. I try to think not only from a few perspectives but from a lot of different perspectives.
12. When I decide something, I try to be objective.
13. When thinking about something, I tend to consider it only from my own perspective.
14. I always try to make unbiased judgments.
15. It concerns me that I might have biases that I am not aware of.
16. When I judge something, I examine the relevant facts and evidence.
17. I do not believe without casting at least some suspicion in every situation.
18. When I conclude, I stick to the concrete evidence that has been presented.

APPENDIX B.

THE CRITICAL THINKING TOOLKIT OR CRITT

Response options:

1 2 3 4 5 6 7 8 9 10

Strongly disagree **Strongly agree**

1. I can detect the use of inappropriate emotional language in scientific arguments.
2. I have a well-defined goal in mind when I am critical.
3. I can identify the structure of arguments without being distracted by their content.
4. Critically thinking is particularly important in psychology.
5. Critical thinking is essential in higher education.
6. When there is a very strong relationship between two variables we can claim that one causes the other.
7. Critical thinking develops as you progress trough your degree.
8. I can express my critical thinking well in my written work.
9. You cannot get a good degree without good critical thinking skills.
10. I prefer to do things where there is a quick answer.
11. I have a focused and systematic way of thinking.
12. All relevant information should be presented in lecture slides.
13. Generally, I am a good critical thinker.
14. I do well in assessments that ask for critical evaluation.
15. I think critically while working on my assignments.
16. All my lecturers expect me to think critically.
17. I know how to approach complex issues in a variety of ways.
18. I will get higher grades if I think critically.
19. I have the ability to judge the value of new information or evidence presented to me.
20. I can evaluate the arguments of others well.
21. Critical thinking is when you describe what is wrong with something.
22. I am good at weighing up both sides of an argument.
23. I can identify analogies between theories.
24. When designing experiments, I can readily eliminate extraneous variables.
25. I think critically while reading.
26. I can rephrase the arguments of others in my own words easily.
27. I think critically in lectures.

APPENDIX C.

PERSONAL WORK STATEMENT

Ondergetekenden / The undersigned:

Karen Bokdam, 122164;

Verklaren ondubbelzinnig dat / Hereby unequivocal declare that:

1) dit werkstuk eigen werk is en derhalve geen inbreuk maakt op het auteursrecht van een ander / The following paper is a personal paper and is therefor no infringement of anyone's copyright.

2) alle gebruikte bronnen (waaronder internetpagina's) zijn voorzien van bronvermelding / All sources used (including websites) have been referred to in footnotes.

3) het verslag voor niet meer dan 5 % aan overgenomen passages uit 'werk van anderen' bevat / The paper is not to include more than 5% of third parties excerpts.

dit verslag ook digitaal is ingeleverd via Blackboard (SafeAssign).

Plaats/Place: Deventer
Datum/Date: 02-06-2020

Handtekeningen/Signature:

Karen Bokdam