Dec 10th, 12:00 AM

Evaluation of IS Curriculum Design: A Pilot Study using the California Critical Thinking Skills Test

Pär J. Ågerfalk
Uppsala University, par.agerfalk@im.uu.se

Jonas Sjostrom
Uppsala University, jonas.sjostrom@im.uu.se

Tuure Tuunanen
University of Oulu, tuure@tuunanen.fi

Evaluation of IS Curriculum Design: A Pilot Study using the California Critical Thinking Skills Test

Short Paper

Pär J. Ågerfalk
Uppsala University
P.O. Box 513, SE-751 20 Uppsala
par.agerfalk@im.uu.se

Jonas Sjöström
Uppsala University
Cramérgatan 3, SE-621 57 Visby
jonas.sjostrom@im.uu.se

Tuure Tuunanen
University of Jyväskylä
P.O. Box 35, FI-40014 Jyväskylä
tuure.t.tuunanen@jyu.fi

Abstract

Critical thinking ability is a key goal of higher education. This pilot study investigates the suitability of using the California Critical Thinking Skills Test (CCTST) to evaluate an IS curriculum design in terms of measuring the students' development of critical thinking skills as they progress through the study program. All currently active students at two similar IS bachelor's programs at Uppsala University was asked to take the test. The results show that the CCTST can be used for the stated purpose. However, the fact that the test is not available in the students' native language is an area of concern. The study also points to the importance of developing strategies for increasing response rate and using relevant indicators of study performance.

Keywords: Curriculum design, evaluation, critical thinking, CCTST

If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts he shall end in certainties.

– Sir Francis Bacon

Introduction

A key concern for any tertiary institution is the quality of their educational programs. In this research, we take inspiration from the Design Science Research (DSR) approach in developing an information systems bachelor's program at Uppsala University in Sweden. Specifically, in keeping with Hevner et al. (2004), we view "organizations, policies, and work practices as designed artifacts" (p. 77) and regard the new study curriculum as a design artifact.

One of the most critical components of DSR is the evaluation of resulting artifacts. Said Hevner et al. (2004, p. 85), "a design artifact must be rigorously demonstrated via well executed evaluation methods". However, there is not yet a clear consensus on how to perform an evaluation in DSR and rigorous evaluation approaches are often hard to apply (Hevner et al. 2004; Venable et al. 2016). Furthermore, despite some fifty years of research on the topic, there is neither consensus on how to evaluate curricula and such evaluation can be done in many different and complementary ways (Welch 1969; White 1971; Leathwood and Phillips 2000; Quek 2017). In this paper, we adopt a quantitative approach to evaluation and focus on one particularly important aspect of a curriculum: the development of students' critical thinking ability.

Critical thinking ability is a key goal of higher education and in Sweden also mandated by the Swedish Higher Education Act (1992:1434). The act states that first-cycle study programs shall develop "the ability of students to make independent and critical assessments" as well as "the ability of students to identify, formulate and solve problems autonomously". Critical thinking includes the capacity to analyse and interpret data, to draw logically correct conclusions in different contexts, and to evaluate and explain various situations. Ultimately, critical thinking is to be content to begin with doubts, as it were. Without a doubt, though, these abilities are central to information systems professionals.

The California Critical Thinking Skills Test (CCTST), a well-established instrument to measure critical thinking skills, has been used to assess college level student's critical thinking ability in a variety of contexts (Al-Fadhli and Khalfan' 2009; Fahim et al. 2012; Pitt et al. 2015). CCTST measures test takers critical thinking skills along seven dimensions (Analysis, Interpretation, Inference, Evaluation, Explanation, Induction, and Deduction). The test is computer-based and time-limited to 45 minutes. Based on the scores on each dimension, an overall score (0–100) is calculated. U.S. based company Insight Assessment administers the test. They collect and analyse the raw data and provide the individual overall score as well as scores per dimension. More information about CCTST is available at https://www.insightassessment.com/.

Except for Al-Fadhli and Khalfan's (2009) evaluation of an e-learning environment, we have not seen the use of CCTST in an IS curriculum evaluation context; and certainly, not within a DSR framework. Accordingly, the purpose of the study at hand is to explore the viability of using CCTST as an instrument for evaluating an IS curriculum design regarding the extent to which its instantiation as a study program develops students' critical thinking skills over time. To this end, we asked currently enrolled students in two similar undergraduate programs at the same university to participate in a pilot study and found a moderate correlation between CCTST score and years of study, which indicates that students indeed develop their critical thinking skills as they progress through the programs. Our main contribution is the validation of CCTST as a viable instrument for evaluating IS curricula. In addition, we provide an example of an approach to rigorous evaluation within a DSR research approach.

In the following, we elaborate in more detail on the adopted DSR based approach to curriculum design. Next, we describe the pilot study and how we intend to use CCTST in a large-scale evaluation of the new curriculum. With this preamble, we report the results of the pilot study and discuss implications for future research on DSR-based curriculum design.

Curriculum Design

The context in which the need for IS curriculum evaluation emerged emanates from the 2013 merger between Uppsala University – the first university in the Nordic countries (est. 1477) – and Gotland University College – the newest tertiary institution in Sweden (est. 1998). Through the merger, the Software Engineering (SE) Department at Gotland was incorporated in the Department of Informatics and Media within the Faculty of Social Sciences at Uppsala University and was assigned the task to develop its two-year undergraduate diploma program in SE into a three-year Computer Information Systems (CIS) bachelor's program.

For the curriculum design, we organized a series of five bi-monthly workshops with teaching staff and representatives from local companies and government agencies. The purpose was to get a better understanding of the needs in the region regarding required graduate skills and to get feedback on the emerging curriculum: the learning outcome, the overall structure, detailed course content, and a plan for industry collaboration. The four primary employers of IS professionals at Gotland, including two government agencies and two companies, participated in all workshops with 2–4 delegates each. The workshops resulted in a series of notes that informed the curriculum design process. Additionally, the design process also included a literature review focusing the last five years of publications on IS and CIS curriculum design. Reflections on early drafts of the emerging curriculum helped formulate design principles (Sjöström et al. 2016) for a DSR based curriculum that introduces students to DSR and design from semester one, as opposed to more traditional curricula where research methods are introduced towards the end of studies. The DSR based curriculum, outlined by Goldkuhl et al (2017), systematically promotes critical thinking by shifting focus between concrete design situations and abstract design reflections, and by promoting evaluation and reflection in connection to all design exercises. Thus, the

learning process is characterized by a continuous reflection about design rationale. The newly designed curriculum also aligns with the Swedish Higher Education Act (1992:1434).

To achieve a rigorous and well-executed evaluation of the new curriculum, we looked for ways to combine established quantitative and qualitative research and assessment methods. The pilot study at hand explores one particular evaluation strategy as a candidate for the main study. The emphasis here is on a quantitative approach that has been used successfully in other domains (Williams et al. 2003; O'Hare and McGuinness 2009; Pitt et al. 2015). The aim of the main study is to conduct a naturalistic evaluation of the design (Venable et al. 2016), i.e. to assess implications for students' learning in a real setting. The pilot study thus concerns the value of CCTST as an instrument to support such evaluation. Ultimately, our ambition is to combine the CCTST approach with qualitative methods to account also for aspects not readily captured quantitatively (Venkatesh et al. 2013; Ågerfalk 2013).

Curriculum Design Evaluation

To evaluate the new curriculum design regarding critical thinking skills development, we plan to implement a longitudinal research design that uses CCTST to measure each student's critical thinking skills (CTS) at three points in time (see Figure 1).

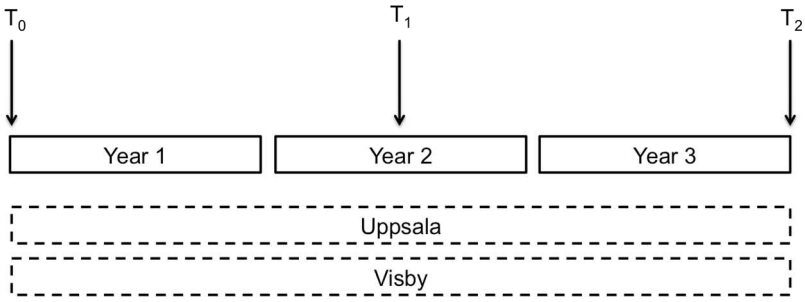

Figure 1. Longitudinal research design to measure students' critical thinking skills development

The first point (T_0) corresponds to the student's entry into the program and thus provides a baseline for future measurements. The second point (T_1) is halfway through the program, and the final point (T_2) is at graduation. Since bachelor's programs in Sweden, following the so-called Bologna model, are three years, T_0 will be at month 1, T_1 at month 18 and T_2 at month 34 (end of year three). To measure progression, we are interested in how each student score at the different measurement points, and in particular their ΔCTS (T_0–T_1, T_1–T_2). Essentially, this is an entry/exit study with an added measurement point in the middle. The additional data will allow us to estimate how much an earlier focus on design research concepts and skills contributes to the students' development of critical thinking skills over time.

In the evaluation, we will leverage the fact that Uppsala University operates two similar programs at its two different campuses. Since the new curriculum will be implemented at the Gotland campus, we plan to use the students at the Uppsala campus for comparison. The program at the Uppsala campus is expected to remain relatively constant throughout the study and any comparatively larger (or smaller) ΔCTS at Campus Gotland can thus, to some extent, be attributed to the new curriculum design.

Pilot Study Design

To validate CCTST as an instrument for our purposes, we first conducted a pilot study with current IS program students in 2016. Our main interest was to find out to what extent CCTST can provide useful data for the longitudinal study design outlined above. Since the pilot study was not longitudinal, we instead recruited respondents from all currently enrolled and active students at all levels of study. More

specifically, we recruited students that had one, two or three years of study by the end of the 2016 spring semester. After that, we conducted a second data collection with newly enrolled students at the start of the autumn 2016 semester. Effectively, this yielded four time points. The pilot study will not provide insights about the progression of individual students but rather about an average score per year for the participating students, which suffices for the purpose of this pilot study.

To establish CCTST as an instrument suitable for our purpose, we sought to achieve high completion ratio and an acceptable response rate. Concerning test results, we expected to find support for the following hypotheses:

h1: Study year → CCTST Score (positive correlation)

h2: CCTST Score → Study performance (positive correlation)

h3: Study background → CCTST Score (positive correlation)

h4: Age → CCTST Score (no correlation)

h5: Gender → CCTST Score (no correlation)

h6: English Skills → CCTST Score (positive correlation)

The first hypothesis (h1: Study year → CCTST Score) captures our intention to measure progression. As suggested by previous research in fields other than IS (O'Hare and McGuinness 2009), the longer one has studied, the higher one would be expected to score on the test. Study year was assigned a value [0, 1, 2, 3] depending on when the student entered into the program.

The second hypothesis (h2: CCTST → Study performance), supported by the findings of O'Hare and McGuinness (2009), suggests that students that score high on the test are likely to be high-performing students in general. To quantify study performance, we used the average number of credit points earned per semester. A full-time student is expected to earn 30 credit points per semester. We used an average score to compensate for the fact that our respondents had been in the program at varying length. A weakness of this measure is that new students have not had time to make any progress, so the number of data points is reduced correspondingly, which makes this construct less reliable.

In keeping with findings from other fields (Vendrely 2007), the third hypothesis (h3: Study background → CCTST) suggests that stronger students are expected to score better on the test. For the main study, we hope to be able to use the Swedish grade point average (GPA) or equivalent. For the pilot, we instead used Campus [Gotland, Uppsala] as a binary proxy since we know that the average GPA is currently lower at the Gotland campus compared to the Uppsala campus.

The two hypotheses (h4: Age → CCTST) and (h5: Gender → CCTST) were intended as controls since the age and gender of the entire population are known. Ages in the group of all current students (as of the 2016 spring semester) ranged from 20 to 44 with an average of 24.33 (median 23, stddev 3.95). In the spring of 2016, 53% of all active students were male and 47% female. It should be noted, though, that previous studies are inconclusive about any significant effect of gender and age on critical thinking skills (Walsh and Hardy 1999).

Finally, self-reported English skills (h6) were used to detect potential language bias. For this purpose, we used the scale suggested by Blue (1993): [Beginner level, Not very good, Adequate, Very good, Native speaker level]. The CCTST is available in many different languages, although not in Swedish, which is the mother tongue of our participants. Fahim et al. (2012) found that critical thinking skills are beneficial for English as a foreign language learners' reading comprehension. Hence, we suspected that there might be an advantage to have a good command of the English language. In fact, it would signal a potential problem with the instrument if this proved to be a strong predictor of CCTST score.

Pilot Study Results

A total of 332 active students were invited to participate in the study (234 in the spring and 98 in the autumn). 34 useable tests were returned, which yields a response rate of just above 10% (0.102). 71% of the test takers were male and 26% female. One individual chose not to reveal their gender. Ages ranged

from 18 to 44 with an average of 25.15 (median 24) and standard deviation of 5.41. 76% of the test takers completed 100% of the test within the allowed 45 minutes, 97% completed at least 71%. The average time spent on the test was 40 minutes (stddev 7.6), and 38% spent the full 45 minutes. The CCTST total score average was 80.68 (stddev 8.13).

We calculated Pearson's correlation coefficients (r), and for comparison, we also calculated Spearman's rank correlations (rho), which turned out to give very similar results. Given the small dataset, these correlations should only be seen as indicative.

We found a moderate correlation between study year and CCTST score (r=0.548, rho=0.561), which tells us that more advanced students scored better. This correlation suggests that the current study programs may help students develop critical thinking skills. The lack of correlation between age and CCTST score (r=0.002) supports this conclusion since had there been such a correlation, life experience could have been a confounding factor.

We were not able to establish a correlation between CCTST score and overall study performance (r=-0.120, rho=0.025). Here we expected to find support for hypothesis h2, i.e. that those students that perform better in class also possess better critical thinking skills and thus score higher on the test. However, because of much fewer data points for this calculation (N=16) as well as the use of average performance per semester as a proxy, one should not put too much weight to this finding. Interestingly, comparing the 16 test takers average number of credit points earned per semester with those of the entire population, the median value is the same for both (30) and the mean value 33.17 (stddev 8.13) and 32.65 (stddev 11.98), respectively. This correspondence suggests that regarding overall study performance, this subset of test takers represents well the overall population, albeit being a somewhat more homogeneous group.

A weak correlation between campus and CCTST score (r=0.296) was found, as expected, and can be explained by the Gotland campus having, for historical reasons, academically weaker students (essentially a lower required GPA to enter into the program). Interestingly, not only the intercept but also the slope of the regression lines differed substantially between the students at Uppsala (y=3.95x+77.79) and Gotland (y=2.87x+74.31). This suggests that the Uppsala students not only start out in a better position but also develop their critical thinking skills at a faster pace than the Gotland students.

A surprising finding was that gender seems to have played a bigger role than anticipated (r=0.241, rho=0.240). This weak correlation shows that the female test takers scored somewhat better than their male counterparts. Compared to the overall student population, females were underrepresented in the study (26% in the study compared to 47% of the overall population). This underrepresentation may indicate a selection bias.

As expected, we found a moderate correlation between (self-assessed) English skills and CCTST score (r=0.514, rho=0.495). Since the CCTST is not available in Swedish (the native language of the students) this could indicate a bias in favour of those students with better knowledge of English. However, a lack of correlation between English skills and Study year (r=0.089, rho=0.070) suggested that these two constructs are independent. To investigate this further we performed a PLS-SEM analysis (Hair et al. 2016), which included the exogenous constructs Study Year and English Skills, the moderator construct English Skills × Study Year and the endogenous construct CCTST score. This analysis showed that there is indeed a weak moderation (β=0.181) in play. The analysis also indicated the influence of Study Year (β=0.495) and English Skills (β=0.501) on CCTST Score.

Discussion and Conclusion

The aim of this pilot study was to validate CCTST as an instrument for evaluating a curriculum design regarding the extent to which its instantiation as a study program develops students' critical thinking skills over time. Four of six hypotheses were supported (h1, h3, h4 and h6), which suggests that CCTST may be a potentially useful instrument for the stated purpose.

Although only a rough estimate due to small sample size, the fact that the slope of the regression line was considerably higher in Uppsala than Gotland is particularly exciting in relation to the main study since the Gotland campus is the one with the redesigned study program and thus where we expect to see improvement.

The moderate correlation between (self-assessed) English skills and CCTST score could, however, indicate a potential problem. Since English skills seem to moderate the influence of study year on CCTST score, this may be a cause of concern and certainly something to pay attention to in the main study.

Overall, we conclude that CCTST worked in the sense that it provided useful data. A major concern, though, is the low response rate (10%). Although prominent IS journals have reported response rates as low as 3% (Sivo et al., 2006), it certainly raises issues of possible non-response error and selection bias. The underrepresentation of females is a related concern. However, a comparison of age and study performance between the respondents and the total population suggests that the respondents are reasonably representative. A lesson learned is that we need to work on a strategy to ascertain participation in the main study. To this effect, we are investigating the possibility of making the test a mandatory part of the study programmes, which could potentially eradicate this issue.

Concerning the two unsupported hypotheses (h2 and h4), the data made us curious to understand better the gender issue and the relationship between CCTST score and overall study performance. Previous research in other disciplinary domains suggests that CCTST score is a significant predictor of study performance (O'Hare and McGuinness 2009) as well as of future job performance (Williams et al. 2003). Understanding better this relationship in an IS context surfaces as an important line of research. Clearly, the average number of credit points earned per semester is a questionable indicator of study performance and more relevant indicators need to be used, such as actual grades and exam scores. Moreover, since CCTST is a general tool for assessing critical thinking skills, it may be beneficial to develop instruments tailored to students of particular disciplines (Pitt et al. 2015). Particularly, future research could therefore develop evaluation instruments tailored to the IS educational context, perhaps inspired by Facione et al.'s (2011) Health Sciences Reasoning Test (HSRT) that has been used extensively in the health sciences (Cox and McLaughlin 2014; Pitt et al. 2015).

CCTST regards critical thinking as a general cognitive ability. A potentially fruitful and complementary strategy may be to view critical thinking as a cognitive function in relation to particular tasks that are relevant to IS professionals (e.g. requirements analysis and conceptual modelling tasks). A productive way to progress in this direction would be to explore the potential of mixed methods, which would allow for addressing both aspects in parallel; adopting a combination of objectivating (critical thinking as a general cognitive ability) and norm-conformative (critical thinking as a cognitive function) attitudes (Ågerfalk 2013). In-depth interviews with students and IS professionals could then provide valuable information about test takers' attitudes towards CCTST and critical thinking skills in general. It may, for instance, well be the case that choosing not to do the test, or perform below one's capability, is a result of one's well developed critical thinking skills. In fact, our plan for the main study is to go beyond the scope of the current paper and include precisely such qualitative elements, including qualitative analyses of final year students' bachelor theses to investigate possible meta-inference (Venkatesh et al. 2013).

References

Ågerfalk, P. J. 2013. "Embracing Diversity through Mixed Methods Research," *European Journal of Information Systems* (22:3), pp. 251-256.

Al-Fadhli, S., and Khalfan, A. 2009. "Developing Critical Thinking in e-Learning Environment: Kuwait University as a Case Study," *Assessment & Evaluation in Higher Education* (34:5), pp. 529-536.

Blue, G. M. 1994. "Self-Assessment of Foreign Language Skills: Does It Work?. *CLE Working Papers*, *3*, 18-35.

Cox, W. C., and McLaughlin, J. E. 2014. "Association of Health Sciences Reasoning Test scores with Academic and Experiential Performance," *American journal of pharmaceutical education* (*78*:4), 73.

Facione, N., Facione, P., Winterhalter, K. 2011. "The Health Sciences Reasoning Test: HSRT — Test Manual," Vol. 2012. Millbrae, CA: The California Academic Press.

Fahim, M., Barjesteh, H., and Vaseghi, R. 2012. "Effects of Critical Thinking Strategy Training on Male/Female EFL Learners' Reading Comprehension," *English Language Teaching* (5:1), 140.

Goldkuhl, G., Ågerfalk, P. J., and Sjöström, J. 2017. "A Design Science Approach to Information Systems Education," In *Proceedings of the 12th International Conference on Design Science Research in Information Systems and Technology*, Springer.

Hair, J. F., Hult, G. T. M., Ringle, C. and Sarstedt, M. 2016 "A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)," SAGE Publications, 2016.

Hevner, A.R., Ram, S., March, S.T., and Park, J. 2004. "Design Science Research in Information Systems," *MIS Quarterly*, (28:1), pp. 75-105.

Leathwood, C., and Phillips, D. 2000. "Developing Curriculum Evaluation Research in Higher Education: Process, Politics and Practicalities," *Higher Education* (40:3), pp. 313-330.

O'Hare, L., and McGuinness, C. 2009. "Measuring Critical Thinking, Intelligence, and Academic Performance in Psychology Undergraduates," *The Irish Journal of Psychology* (*30*:3-4), pp. 123-131.

Pitt, V., Powis, D., Levett-Jones, T., and Hunter, S. 2015. "The Influence of Critical Thinking Skills on Performance and Progression in a Pre-Registration Nursing Program," *Nurse education today* (35:1), pp. 125-131.

Quek, C. G. 2017. "Curriculum Evaluation," In *Curriculum for High Ability Learners: Issues, Trends and Practices*, L. S. Tan, L. D. Ponnusamy, and C. G. Quek (Eds.), Singapore: Springer Singapore, pp. 223-239.

Sivo, S. A., Saunders, C., Chang, Q., and Jiang, J. J. 2006. "How Low Should You Go? Low Response Rates and the Validity of Inference in IS Questionnaire Research," *Journal of the Association for Information Systems* (7:6), pp. 35-414.

Sjöström J., Ågerfalk P. J., and Tunnanen, T. 2016. "Five Principles for DSR Based Curriculum Development," in Proceedings of the AIS SIGED 2016 Conference, Dublin, Ireland, December 2016.

Venable, J., Pries-Heje, J., and Baskerville, R. 2016. "FEDS: A Framework for Evaluation in Design Science Research," *European Journal of Information Systems* (25:1), pp. 77-89.

Vendrely, A. M. 2007. "An Investigation of the Relationships among Academic Performance, Clinical Performance, Critical Thinking, and Success on the Physical Therapy Licensure Examination," *Journal of allied health* (36:2), pp. 108-123.

*V*enkatesh, V., Brown, S. A., and Bala, H. 2013. "Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems," *MIS Quarterly* (37:1), pp. 21-54.

Walsh, C. M., and Hardy, R. C. 1999. "Dispositional Differences in Critical Thinking Related to Gender and Academic Major," *Journal of Nursing Education* (38:4), pp. 149-155.

Welch, W. W. 1969. "4: Curriculum Evaluation," *Review of Educational Research* (39:4), pp. 429-443.

White, J. P. 1971. "The Concept of Curriculum Evaluation," *Journal of Curriculum Studies* (3:2), pp. 101-112.

Williams, K. B., Glasnapp, D. R., Tilliss, T. S., Osborn, J., Wilkins, K., Mitchell, S., Kershbaum, W., and Schmidt, C. 2003. "Predictive Validity of Critical Thinking Skills for Initial Clinical Dental Hygiene Performance," *Journal of dental education* (67:11), pp. 1180-1192.