

UNIVERSITÄT
DUISBURG
ESSEN

Open-Minded

DOCTORAL THESIS

Supportive guidance methods for wiki-based learning and knowledge construction

Faculty of Engineering,
Department of Computer Science and Applied Cognitive Science
at the University of Duisburg-Essen

*A thesis approved in fulfilment of the requirements for the degree of
Dr. rer. nat.*

by

Sven HEIMBUCH

from Bochum

Examined by

Prof. Dr. Daniel BODEMER

Prof. Dr. Ulrike CRESS

Date of oral examination: March 9, 2018

“It would seem that you have no useful skill or talent whatsoever,” he said. “Have you thought of going into teaching?”

– Terry Pratchett, *Mort*

Acknowledgements

Completing this dissertation was certainly one of my biggest challenges. It would be a blatant lie if I said this journey was always clear and unambiguous. Many times, I had doubts in myself. But there were some people who repeatedly brought me back on track, and I would like to thank them! I would like to thank my supervisor, Daniel Bodemer. In the beginning you were my superior, who is still my boss at this time, but you also became a friend to me. You helped me with advice, not only in questions of theory and methodology, but also had an open ear for personal matters. This was the only way for me to take care of the other side of my life, my beloved son Felix. You and my significant other Hanna have regularly brought me back to reality and down to earth, reminding me that PhD and work aren't everything. Playing is sometimes the only thing that matters! I can't thank you both enough for all you have done for me. I would also like to thank my colleagues Marielle Dado, Melanie Erkens and Lenka Schnaubert for all the talks and discussions we have had on all kinds of things. I must emphasize the many opportunities of productive procrastination with Lenka Schnaubert. It has always been a great pleasure for me to turn your methodological puzzles and problems into my problems for a short time. I learned a lot and it was always a good distraction to get focused again afterwards. Finally, I would like to thank the following students: Lisa Ollesch, Kerstin Uhde and Holger Wallstab. I could advise you during the studies of your bachelor's theses and to help you with words and deeds. And the data you collected in these studies helped me to complete this PhD thesis.

Declaration of Authorship

I, Sven HEIMBUCH, declare that this thesis titled, "Supportive guidance methods for wiki-based learning and knowledge construction" and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a research degree at the University of Duisburg-Essen.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- The work was not submitted to any other review committee in the same or similar form.

Signed: 

Date: March 26, 2018

Publication Notes

Parts of this thesis were presented at conferences or published in peer-reviewed journals. Please refer to the following publications for more information:

Experiment 1

Heimbuch, S. & Bodemer, D. (2014). Supporting awareness of content-related controversies in a wiki-based learning environment. In *Proceedings of The International Symposium on Open Collaboration* (30:1–4). New York, NY, USA: ACM. <https://10.1145/2641580.2641607>.

Heimbuch, S. & Bodemer, D. (2016a). Effects of implicit guidance on contribution quality in a wiki-based learning environment. In *Proceedings of the 12th International Conference of the Learning Sciences (ICLS 2016)* (Vol. 2, pp. 906–909). International Society of the Learning Sciences.

Heimbuch, S. & Bodemer, D. (2017). Controversy awareness on evidence-led discussions as guidance for students in wiki-based learning. *The Internet and Higher Education*, 33(2017), 1–14. <https://10.1016/j.iheduc.2016.12.001>.

Experiment 2

Heimbuch, S. & Bodemer, D. (2015a). Let's Talk about Talks: Supporting Knowledge Exchange Processes on Wiki Discussion Pages. In *Ninth International AAAI Conference on Weblogs and Social Media (ICWSM-15)* (pp. 56–61). AAAI Press.

Experiment 3

Heimbuch, S., Uhde, K., & Bodemer, D. (2014). "a new Wiki way?" - An experimental study of collaborative knowledge building scripts. In A. F. Mohd Ayub, T. Matsui, C.-C. Liu, H. Ogata, S. C. Kong, & A. Kashihara (Eds.), *Proceedings of the 22nd International Conference on Computers in Education (ICCE 2014)* (Vol. WIPP, pp. 10–12). Nara, Japan: Asia-Pacific Society for Computers in Education.

Experiment 4

Heimbuch, S., Ollesch, L., & Bodemer, D. (2016). Effects of collaboration scripts on learning activities in wikis. EARLI SIG 7: Learning and Instruction with Computers. Dijon, France: European Association for Research on Learning and Instruction.

Experiment 5

Heimbuch, S. & Bodemer, D. (2015b). Need for cognitive closure as determinant for guidance in Wiki-based learning. In T. Matsui, A. F. Mohd Ayub, B. Jiang, H. Ogata, W. Chen, S. C. Kong, & F. Qiu (Eds.), *Proceedings of the 23rd International Conference on Computers in Education (ICCE 2015)* (Vol. WIPP, pp. 10–12). Hangzhou, China: Asia-Pacific Society for Computers in Education.

Heimbuch, S. & Bodemer, D. (2016b). Wiki Editors' Acceptance of Additional Guidance on Talk Pages. In *Tenth International AAAI Conference on Web and Social Media (ICWSM-16)* (pp. 51–52). AAAI Press.

Zusammenfassung

Wikis gelten als besondere Vertreter sozio-technischer Systeme, die zunehmend für die kollaborative Konstruktion von Wissen und darüber hinaus für individuelles und kollaboratives Lernen genutzt werden. Das grundlegende Design von Wiki-Systemen ermöglicht es den Nutzern, Inhalte als Artikel zu generieren und auf zugehörigen Diskussionsseiten ("Talk Pages") im Artikelhintergrund über Themen zu diskutieren. Aufbauend auf bestehenden Theorien und gegenwärtiger Forschung zum Wissensaufbau mit Wikis, und darüber hinaus aus dem Bereich des computergestützten kollaborativen Lernens (CSCL), wurden in dieser Dissertation mehrere Effekte auf Prozesse und Ergebnisse untersucht in Bezug auf zusätzliche ergänzende Strukturierungsmaßnahmen für wiki-basiertes Lernen. Konkret konzentrierte sich diese Arbeit auf (1) die Effektivität und Effizienz impliziter Lenkung für Wiki-Talk-Seiten, (2) die Auswirkungen zweier Ansätze von Kollaborationsskripten als explizite Lenkung zur Wissenskonstruktion mit Wikis und (3) die Relevanz spezifischer lernbezogener individueller Unterschiede für kollaboratives Lernen mit Wikis. Insgesamt wurden im Rahmen dieser Dissertation fünf empirische Studien durchgeführt. Mit Studie 1 wurden die Auswirkungen von visuellen Hervorhebungen kontroverser Argumentationen für Wiki-Diskussionen untersucht. Die Ergebnisse des Experiments zeigten, dass zusätzliche Hervorhebungen für kontroverse Diskussionen sich direkt auf das individuelle Selektions- und Leseverhalten sowie indirekt und in geringerem Maße auf die Lernergebnisse und die Qualität von Wiki-Beiträgen

auswirken. In Studie 2 wurde untersucht, ob Visualisierungen von Autorenwissen und Beurteilungen durch die Wiki-Community implizit die Wahrnehmung die Leser von kontroversen Diskussionen in Wikis beeinflussen. Die Ergebnisse zeigten, dass, wenn zusätzliche Autoreninformationen visualisiert werden, es wahrscheinlicher ist, dass Leser von Wiki-Diskussionen der Argumentation eines angenommenen Experten folgen. In den Studien 3 und 4 wurden die Effekte zweier unterschiedlicher Kollaborationsskript-Ansätze untersucht. Das erste Skript wurde von Wikipedia abgeleitet, während das zweite Skript ein selbst entwickeltes Skript ist, das von empirischen Untersuchungen inspiriert wurde. Die Ergebnisse zeigten, dass der alternative Skriptvorschlag vorteilhafter ist in Bezug auf Perspektivübernahme und Integration widersprüchlicher Belege, sowie für den individuellen Lernerfolg und die Qualität kollaborativ bearbeiteter Artikel. Studie 5 untersuchte die Auswirkungen der Hervorhebungen für kontroverse Diskussionen und des alternativen Kollaborationsskripts im Zusammenspiel mit individuellen Unterschieden des kognitiven Geschlossenheitsbedürfnisses. Dieses Konstrukt ist relevant für das Verständnis, wie Menschen zweideutige Informationen verarbeiten, die in kontroversen Diskussionen zu finden sind. Die Ergebnisse zeigten, dass Personen mit einem hohen Bedürfnis nach kognitiver Geschlossenheit in Bezug auf den Lernerfolg mehr von den Hervorhebungen für kontroverse Diskussionen als implizite Lenkung profitieren, während Personen mit einem niedrigen Bedürfnis mehr von dem Kollaborationsskript als explizite Lenkung profitieren. Diese Studienreihe erweitert die empirische Basis der Forschung zu wiki-basierter Wissenskonstruktion und Lernprozessen um Untersuchungen zu ergänzenden unterschiedlichen Strukturierungsmaßnahmen und die Berücksichtigung individueller Unterschiede.

Abstract

Wikis are a special representative of socio-technical systems that are increasingly used for the collaborative construction of knowledge and furthermore for individual and collaborative learning. The basic design of wiki systems enables users to generate content as articles and as well to discuss about subject matters on corresponding discussion forums in the article background, the so-called talk pages. Building upon prevailing theories and previous research on knowledge building with wikis, and more broadly computer supported collaborative learning (CSCL) in general, this dissertation investigated several effects of added supplemental scaffolding measures for wiki-based learning on processes and outcomes. Specifically, this work focused on (1) the effectiveness and efficiency implicit guidance approaches for wiki talk pages, (2) the effects of two distinct collaboration scripts as explicit guidance for knowledge construction with wikis, and (3) the relevance of specific learning-related individual differences for collaborative learning with wikis. Overall, five empirical studies have been conducted as part of this dissertation. Study 1 examined the effects of added controversy awareness highlights for wiki discussions. Results of the experiment showed that added highlights for controversial discussions directly affect individual selection and reading behaviour, as well as indirectly and to a lesser extent the learning outcomes and wiki contribution quality. Study 2 examined whether visualisations of author expertise and community-rating implicitly affect the user perception of controversial discussions in wikis. Results showed that if additional author information is visualised, it is much more likely that readers of

wiki discussions follow an assumed expert's argumentation. Studies 3 and 4 both examined effects of two distinct collaboration script approaches. The first script was derived from Wikipedia, whereas the second script is a self-developed script that was inspired by related empirical research. Results showed that the alternative script proposal is more beneficial for perspective-taking and integration of opposing evidence, as well as for individual learning success and the quality of collaboratively edited articles. Study 5 examined the effects of the controversy awareness highlights and the alternative collaboration script in interaction with individual differences of the Need for Cognitive Closure. This construct is relevant for the understanding of how people process ambiguous information that are likely to be found in controversial discussions. Results showed that persons with a high Need for Cognitive Closure benefit more, in terms of learning success, from the controversy awareness highlights for implicit guidance, whereas persons with a low Need for Cognitive Closure benefit more from the collaboration script as explicit guidance. This study series extends the empirical base of research on wiki-based knowledge construction and learning processes with investigations of supplemental different guidance measures and the consideration of individual differences.

Contents

Acknowledgements	IV
Declaration of Authorship	V
Publication Notes	VI
Zusammenfassung	VIII
Abstract	X
I Theoretical and Empirical Backgrounds	1
1 Introduction and Overview	3
1.1 Collaborative content creation in wikis	8
1.2 Processes of (collaborative) writing	10
1.3 Knowledge building and wikis	13
1.4 The co-evolution of knowledge	14
1.5 Socio-cognitive conflicts	17
1.6 Need for help in wiki-based learning	19
1.6.1 Helping to judge an information source	22
1.6.2 Influences of individual differences	26
1.7 Guiding and supporting learners	35
1.7.1 Implicit guidance	35
1.7.2 Explicit guidance	37
1.8 General Research Questions	41
1.9 Remarks on Statistical Methods	49

II	Experimental Studies	53
2	Controversy Awareness on Evidence-led Discussions as Guidance for Students in Wiki-based Learning	55
2.1	Method	56
2.1.1	Design and Participants	56
2.1.2	Material	58
2.1.3	Wiki environment	59
2.1.4	Measurements	61
2.1.5	Procedure	64
2.1.6	Specific Hypotheses to Research Questions	66
2.2	Results	68
2.3	Discussion	84
2.3.1	Outlook	91
3	Representations of Authors' Community-Standing and Expertise Information on Wiki Talk Pages	93
3.1	Method	94
3.1.1	Design and Participants	94
3.1.2	Materials	94
3.1.3	Measurements	96
3.1.4	Study Procedure	98
3.1.5	Specific Hypotheses to Research Questions	99
3.2	Results	100
3.3	Discussion	106
4	Investigating Collaboration Scripts for Learning with Wiki-Based Environments	111
4.1	General Method of Experiments 3 & 4	112
4.1.1	Specific Hypotheses to Research Questions	112
4.2	Method of Experiment 3	114

4.2.1	Design and Participants	114
4.2.2	Procedure and Materials	115
4.2.3	Variables and Measurements	117
4.3	Results of Experiment 3	119
4.3.1	Direct Group Comparisons	119
4.3.2	Influencing Variables	122
4.3.3	Summary of Main Findings	123
4.4	Method of Experiment 4	124
4.4.1	Participants and Design	124
4.4.2	Materials and Procedure	125
4.4.3	Variables and Measurements	128
4.5	Results of Experiment 4	129
4.5.1	Analytic Procedure	129
4.5.2	Direct Group Comparisons	130
4.5.3	Conditional Effects on Contribution Frequency	131
4.5.4	Conditional Effects on Learning	133
4.5.5	Mediation of Contributions Between Script and Learning	134
4.5.6	Further Results	135
4.5.7	Summary of Main Findings	136
4.6	Discussion	137
5	The Interaction of the Need for Cognitive Closure with Implicit and Ex-	
	PLICIT GUIDANCE IN WIKI-BASED LEARNING	141
5.1	Method	142
5.1.1	Design and Participants	142
5.1.2	Procedure	143
5.1.3	Materials and Wiki Environments	145
5.1.4	Variables and Measurements	147
5.1.5	Specific Hypotheses to Research Questions	150

5.2	Results	151
5.2.1	Testing Equivalence between Guidance Groups	152
5.2.2	Testing Equivalence between NCC Levels	154
5.2.3	Testing Interactions of Guidance and NCC	156
5.2.4	Further Measurements	159
5.3	Discussion	160
III	General Discussion	165
6	General Discussion	167
6.1	Summary of Findings	168
6.2	Theoretical Implications	172
6.2.1	Implications for the co-evolution model	172
6.2.2	Implications for implicit guidance	174
6.2.3	Implications for explicit guidance	175
6.2.4	Further implications	176
6.3	Practical Implications	177
6.4	Outlook and Future Directions	179
6.5	Final Conclusion	180
	References	182

List of Figures

1.1	Cognitive process model of writing. Adapted from Flower and Hayes (1981)	11
1.2	Schematic representation of the occurrence of a socio-cognitive conflict, in line with the co-evolution model. An individual with prior knowledge about Evidence A perceives a controversy between discussants in the social system with arguments for Evidence A and Evidence B. The reception of Evidence B in the social system conflicts with the individual's prior knowledge about Evidence A.	19
1.3	Framework of the help-seeking process, adapted from Makara and Karabenick (2013). The stages highlighted in bold represent starting points of guidance measures for wikis as discussed in Experiments 1 and 2 of this thesis (cf. Chapter 2 and 3, respectively).	20
1.4	In this model representation, metacognitions (meta-level) and cognitions (object-level) are interrelated with each other, sharing information with each other by means of monitoring and control processes. Figure adapted from Nelson and Narens (1990).	33
1.5	Representations of the Be Bold, Revert, Discuss (left) and Discuss, Deliberate, Revise (right) scripts main steps as used in Experiments 3 and 4 (cf. Chapter4) as a flow chart. In Experiment 5 only the Discuss, Deliberate, Revise was used cf. Chapter5.	40

1.6	In this model hypotheses of the major research questions of the experiments that were conducted for this thesis are represented. Hypotheses encompassed direct and indirect (via mediation and moderation) influences of additional guidance for wikis on process and outcome variables.	48
1.7	Mean differences (black squares) and 90% confidence intervals (horizontal lines) with equivalence bounds $\Delta L = -0.5$ and $\Delta U = 0.5$ for four combinations of test results that are statistically equivalent or not, and statistically different from zero or not. Pattern <i>A</i> is statistically equivalent, pattern <i>B</i> is statistically different from 0, pattern <i>C</i> is practically insignificant, and pattern <i>D</i> is inconclusive (neither statistically different from 0 nor equivalent). Figure adapted from Lakens (2017a). .	51
2.1	Excerpt of the study's generated talk page, depicting a discussion thread comprising a resolved evidence-led controversy between two discussants in one of the experimental groups.	56
2.2	Excerpt of the study's generated talk page, depicting a discussion thread comprising a resolved evidence-led controversy between two discussants in one of the experimental groups.	59
2.3	Illustration of a talk page excerpt for all groups: no highlight (top) vs. controversy highlight (middle) vs. controversy + status highlight (bottom).	60
2.4	Workflow diagram visualising the overall study procedure of Experiment 1 with its central stages.	66
2.5	Markov models with state transition probabilities for <i>no highlight (control)</i> (left), <i>controversy highlight (middle)</i> and <i>+ status highlight (right)</i> . .	70

2.6	Distributions of absolute reading times in seconds for the discussion categories.	72
2.7	Path analysis on reply quality with process variables (selecting, reading, replying) as serial mediators. Solid lines represent regressions below an $\alpha = .05$ threshold, with bolder lines representing lower p -values.	76
2.8	Path analysis on article edit quality with process variables (selecting, reading, replying) as serial mediators. Solid lines represent regressions below an $\alpha = .05$ threshold, with bolder lines representing lower p -values.	78
2.9	Multiple single-step mediation model on the multiple-choice test results as an outcome variable. Unstandardised a and c' weights are measured in seconds. Unstandardised b weights are measured in number of correct answers. ** $p < .01$, *** $p < .001$	80
2.10	Multivariate interaction graphs between groups and the individual Need for Cognitive Closure on replies within the study's different discussion types (resolved vs. unresolved vs. peripheral).	81
3.1	Sample page excerpt from a wiki talk page discussion with additional author information. Ranking badges reflect domain-specific expertise ratings of authors and W+ (Wikes) reflect the overall wiki author reputation regardless of the domain as rated by the community.	96
3.2	Workflow diagram visualising the overall study procedure of Experiment 2 with its central stages.	99
3.3	Conditional effects of the Need for Cognitive Closure on discussion reading times. Dashed lines = default wiki group, solid lines = experimental wiki group.	102

3.4	Johnson-Neyman (JN) conditional effect of social recommendations on perceived clarity of the talk page discussions moderated by the individual Need for Cognitive Closure. Dashed lines represent the upper limit (ULCI) and lower limit (LLCI) of the conditional effect's 95% confidence interval. The grey shaded area represents the JN region of significance. It is bound between the moderator's (Need for Cognitive Closure) relative values of 0.25 and 0.82 (dotted vertical lines) and below the zero line (dotted horizontal line).	103
3.5	Scale results of the six dimensions of the user experience questionnaire UEQ for the additional information (ranking badges and Wikes) as used in the experiment's modified wiki.	106
4.1	Hypothesis model representing which of the experiments (3 and 4) addressed what effects, dependent variables and influencing variables.	113
4.2	Wiki-mimicking learning environment with collaboration script representation in Experiment 3.	116
4.3	Workflow diagram visualising the overall study procedure of Experiment 3 with its central stages.	117
4.4	Incorporation of viewpoints of one's own learning material and the learning partner's material at t1 (left) vs t2 (right).	121
4.5	Matrices of need for cognitive closure scatterplots for Be Bold, Revert, Discuss script (top panel) and Discuss, Deliberate, Revise script (bottom panel).	123
4.6	Workflow diagram visualising the overall study procedure of Experiment 4 with its central stages.	126

4.7	Screenshot of one experimental group's wiki-based learning environment with collaboration script representations as used in Experiment 4.	127
4.8	Conditional effects of script group on contributions as functions of influencing variables: need for cognitive closure (left), metacognitive strategies (middle), intrinsic motivation (right). Grey area = JN region of significance, dashed lines = 95% upper and lower confidence bands.	132
4.9	Simple slopes of script groups on learning gains conditional of influencing variables: need for cognitive closure (left), metacognitive strategies (middle), intrinsic motivation (right). Dashed lines = BRD script, solid lines = DDR script.	134
4.10	Effect of collaboration script on learning outcome mediated by the frequency of contributions to the wikis.	135
5.1	Workflow diagram visualising the overall study procedure of Experiment 5 with its central stages.	144
5.2	The implicit guidance wiki with controversy highlights is represented in the floating upper left part of the figure. A DDR collaboration script representation for the explicit guidance wiki is shown on the right.	146
5.3	Equivalence hypothesis test using the two one-sided <i>t</i> -Tests (TOST) procedure between experimental groups on mean test scores.	153
5.4	Equivalence hypothesis test using the two one-sided <i>t</i> -Tests (TOST) procedure between low and high NCC levels on mean test scores.	155

5.5	Interaction diagrams of the guidance type with Need for Cognitive Closure (NCC) on knowledge test scores. First, the interaction is visualised as in a 2x2 ANOVA with the dichotomised (median split) NCC (top panel). Second, the linear regression slopes are visualised with smoothed standard error areas for the continuous NCC spectrum (bottom panel).	158
5.6	Mean ratings on the User Experience Questionnaire dimensions for controversy awareness highlights for implicit guidance implicit (blue) and collaboration script DDR for explicit guidance (red) implementations.	160
6.1	The direct and indirect influences of additional guidance measures on processes and outcomes. Light green represents moderate evidence for effects. Dark green represents stronger evidence for effects. The overall evidence for the constructs and variables is cumulative over all conducted studies.	168
6.2	Proposal of a unified model for collaborative knowledge construction with wikis. Implicit guidance can be used to influence the help-seeking process at different stages. Going through this process can lead to contributions in the wiki that are determined by interchanging processes of internalisation and externalisation which can be influenced by explicit guidance. The effectiveness and efficiency of both types of guidance can be influenced by individual differences.	176

List of Tables

1.1	A classification scheme for Bayes Factors as suggested by Lee and Wagenmakers (2013).	52
2.1	Example questions/statements and answers of the 15-item multiple-choice test.	62
2.2	Most frequent closed sequential patterns discovered by the CM-ClaSP algorithm.	69
2.3	In-depth analyses on the relative reading times of the different discussion thread categories.	71
2.4	Multivariate Helmert contrast comparisons on replying to discussion categories.	74
2.5	Frequencies (and percentages) of cumulative overall quality ratings of discussion replies.	74
2.6	Frequencies (and percentages) of integrating summaries of the preceding discussion into replies.	75
2.7	Frequencies (and percentages) of cumulative overall quality ratings of edited articles.	77
2.8	In-depth analyses on the number of correct answers in the knowledge test.	79
2.9	Correlations between epistemic curiosity and topic selections.	83
2.10	Correlations between epistemic curiosity and topic reading times.	83
3.1	Welch's <i>t</i> -Tests on correct controversy categorisation.	104

4.1 Titles and lengths of the original wiki articles that were created in the
previous semester. 125

This thesis is dedicated to my
son Felix Sami who constantly
reminds me what really matters
in life.

Part I

Theoretical and Empirical

Backgrounds

1 Introduction and Overview

Ward Cunningham, the programmer of the first wiki, originally described wikis as “the simplest online database that could possibly work” (Cunningham, 2002). The most significant representative of a wiki is undoubtedly the free online encyclopaedia Wikipedia. It serves as a platform for the collection and collaborative construction of knowledge. Wikis are web-based platforms that enable collaboration at virtually any point in time from any place between any person having access to a web-enabled device (Chen, Jang, & Chen, 2015). They allow one or more persons to build up a corpus of knowledge in a set of interlinked web pages using a process of creating and editing pages (Franklin & Harmelen, 2007). Knowledge can be shared at any time with a whole community without major restrictions, while the focus of this thesis lays on the learning aspects rather than on the aspect of information sharing. This is inspired by the constructivist learning theory which assumes that individuals learn best when they have constructed knowledge on their own (Cole, 2009). Consequently, wiki environments are nowadays widely used in many educational contexts for collaborative knowledge construction and learning tasks (Notari, Reynolds, Chu, & Honegger, 2016). Wiki article talk pages are a special kind of discussion forum that go along with the articles and differ by a considerable extent from classic online discussion forums. Each talk page is like every other page in a wiki editable by every user. Replies to a certain contribution are not appended automatically as in threaded discussion forums, they must be integrated within the wiki structure by the user. It is also possible to change the content of the contributions by

others, although (semi-)automated measures to prevent vandalism are prevalent in many wiki systems. It is like with regular threaded discussion forums that authors and their contributions are directly identifiable by their usernames on talk pages, in contrast to what is visible in a wiki article. This kind of rather special discussion forums can provide a fertile ground for controversies and socio-cognitive conflicts to occur, either between discussants with heterogeneous knowledge and opposing arguments, or between a reader's observations of controversial discussions and their own contradictory prior knowledge and viewpoints. Wikis such as Wikipedia or Wikiversity give a common and widespread opportunity to share user generated content that emerges from collaborative writing processes. These contents are either available publicly over the Internet or within closed networks such as universities that supply wikis for study coursework or learning groups. Prevailing cognitive theories of writing and knowledge construction of socially shared artefacts help to identify, understand and effectively use individually different prerequisites of wiki users. Researchers and instructors should consider several individual differences before writers can be enabled to effectively collaborate with others in a collaborative knowledge construction environment, since collaboration just does not simply occur because of the use of a collaborative environment (Notari et al., 2016).

A potential measure to create an added benefit for wiki users is making individual wiki readers and editors aware of controversial perspectives within discussed topics that have differing statuses of resolution. This will be addressed in this thesis in Experiment 1 that will be presented and discussed in Chapter 2. There, an experimental study with university students is presented that investigates potential guidance effects of visual controversy awareness support on discussion threads and analyses of underlying learning mechanisms from an individual learner's perspective. If taking Wikipedia as a prime example, over its entire system space an extremely

high number of discussions about corresponding articles exists (Wikipedia, 2017c). It has more than 5.4 million articles in the English language version alone, with high-quality articles that were generated in a relatively brief period compared to traditional encyclopaedias (Wikipedia, 2017c). For instance, the WikiMedia Foundation implemented a system of community-based quality reviews in Wikipedia, such as assigning the label “featured content” to excellent articles (Wikipedia, 2017b). However, this label reflects exclusively the overall quality of collaboratively constructed user-generated content. It does not refer to the quality of an individual contribution to an article or the related discussion that took place in the background of an article, nor does it reflect the expertise or credibility of an author. In Chapter 3 Experiment 2 will be discussed which addresses potential effects of including social recommendation information about wiki authors on talk page discussions. Despite the common and wide dissemination of computer-supported collaborative learning (CSCL) environments such as wikis, the effective and meaningful use of them can still be quite challenging for learners (Zheng, Niiya, & Warschauer, 2015). Some of these challenges for collaboration with wikis can be the abundance of options and unfamiliar structure. That can make added guidance crucial, because unguided collaboration does not automatically result in learning benefits for users of such systems (Dillenbourg, 2002). Thus, one way of supporting individuals in computer-based collaborative learning environments are explicit instructions to organise their interactions (Bromme, Hesse, & Spada, 2005). Instructional scaffolding as added guidance is often needed, so that the full potentials of collaborative learning can be exploited. Without such guidance, the full potential of wikis with its multiple layers may not be used, and they are only used as environments for collaborative article production and development without deeper elaborations (Biasutti, 2017). To

analyse which type of explicit guidance is most beneficial for learners in wiki environments, two experimental studies will be discussed in Chapter 4 where students had to collaboratively construct knowledge artefacts either in dyads (Experiment 3) or in larger groups (Experiment 4).

Although wikis in academia are nowadays widely acceptance, their effectiveness and efficiency for collaborative learning activities are inconclusive due to ambiguous results. Research on wikis has proposed added scaffolding measures to be incorporated in wikis to improve the overall quality of writing and coordination processes (Papadopoulos, Demetriadis, & Weinberger, 2013; Wichmann & Rummel, 2013). Socio-technical environments like wikis offer many opportunities to discuss any article's contents on their corresponding talk pages. Essential to the co-creation of user-generated content are processes of internalizing and externalizing knowledge from an individual into the wiki or vice versa (Cress & Kimmerle, 2008). Within these processes lay potentials for emerging controversies between collaborators that are grounded on different opinions or contradictory knowledge. These controversies can induce socio-cognitive conflicts that trigger individual equilibration and elaboration processes which can be beneficial for learning (Mugny & Doise, 1978). In earlier research, essentially two types of scaffolds have been successfully evaluated both focusing on potential benefits for talk pages: (1) visual representations as implicit guidance (Heimbuch & Bodemer, 2014, 2017), (2) collaboration scripts as explicit guidance (Heimbuch & Bodemer, 2015a; Heimbuch et al., 2014). The former is meant as a minimal invasive modification to make controversial discussions and their concurrent state more salient by adding visual highlights. Whereas the latter is proposed to be an alternative to Wikipedia's *Be Bold* principle and aims at fostering deeper elaboration processes by encouraging discussions prior to the externalization of knowledge into a wiki. Both types of proposed scaffolds for wikis

aim at focusing readers and editors towards relevant contents on discussion pages and by extension provide guidance for content creation and knowledge construction processes. The present study compares participants' adoption of both types of guidance support in two experimental wiki-based learning environments and analyses the individual effects on wiki-related activities. In this thesis, potential benefits will be discussed when added guidance is supplied to users of wiki talk pages to make information about controversial discussion and author expertise salient and also to guide them into the process of discussing planned article changes. When dealing with controversies and socio-cognitive conflicts that eventually arise out of this, it can be crucial to focus on individual differences that address the processing of ambiguity in information. There are indications that the effort a learner is willing to invest in searching for solutions to a problem can be influenced by the individual's Need for Cognitive Closure. This cognitive construct stands for a motivational continuum between the need to get a clear answer in an ambiguous situation and the avoidance of quick and unambiguous answers. Various empirical results and discussions illustrate that it can be regarded as a relevant construct in knowledge creation processes (Webster & Kruglanski, 1994; Heimbuch & Bodemer, 2015b). Although there are close ties between the Need for Cognitive Closure and inter-individual differences in learning and knowledge construction, there are only few studies in technology-enhanced learning to address this construct. The interaction of effects between the kind of provided guidance for wiki users with the individual Need for Cognitive Closure will be discussed with an experimental study (Experiment 5) in Chapter 5. In the following subsections, the previously in-brief outlined theories and empirical research will be discussed in more detail.

1.1 Collaborative content creation in wikis

Wikis are collaborative platforms where everyone can add more additional content and has also the ability to update or edit already existing content (Richardson, 2006). One of the main objectives of wikis is to share knowledge with a community without levying larger restrictions at any time. Furthermore, wikis as web-based tools extensively document any changes made to the environment and thus give permanently the opportunity for valuable insights into processes of collaborative knowledge construction (Biasutti & EL-Deghaidy, 2012). However, evidence-led controversies and their underlying arguments can enable multiple perspectives on a topic and thus be very beneficial. An interested reader can broaden her or his own knowledge. As a potential new contributor, a user could also provide new knowledge and insights into a discourse. Other kind of controversies that are observable in wiki-environments can be on a formal level about the general structures of articles or on a social interpersonal level between users (Kittur & Kraut, 2010). As such, wikis have already been used in various educational settings and have been integrated in many applications and assignments used for teaching (Bartelsen & Brauer, 2010). Especially in higher education, they can be implemented in almost all kinds of degree course programmes, to facilitate collaborative learning of new definitions and concepts. In addition to a relatively intuitive operation of the basic functions trying to simplify collaborative learning and working, the underlying structure and mechanisms of wikis can help with learning processes and social interactions, thereby offering many opportunities for collaborative knowledge construction. From a constructivist's perspective, wikis inherently have great potentials for collaborative learning where people learn best when they designed their learning material by themselves (Cole, 2009).

User-generated content in wikis like Wikipedia is mainly generated on two distinct layers of the environment. Built into the social system's core foundations, there is an explicit division between an article page and its corresponding talk page consisting of article-related discussion. Wikipedia talk pages include many exchanges of opposing controversial evidence and arguments on topics that are often not reflected in an article, due to the Wikipedia's neutral point of view (NPOV) guideline (Wikipedia, 2016). This guideline makes sure that most content is written neutrally with no obvious controversies in the main article body. If controversies about a certain topic exist, NPOV requires that they must be clearly identifiable, e.g. under a dedicated subheading. As a further very fundamental wiki principle, in most cases there are no exclusion criteria for participation in any activity related to creating and changing content. In other words, every user has virtually equal rights in the social systems. Thus, there is the freedom to edit one's own created content, a third party's content or discuss about uncertainties and give feedback to changes performed by others (Chen et al., 2015). To date, much wiki-related research has predominantly investigated the article content quality or effects of direct interventions at the article level as the primary layer of wiki interaction (Agasta Adline & Mahalakshmi, 2012; Lucassen & Schraagen, 2011; Reiterer, Stettinger, Jeran, Eixelsberger, & Wundara, 2015; Wichmann & Rummel, 2013). In other related research extensive log data analyses were performed on directly accessible wiki metrics, such as the article revision history with the aim of identifying article editing patterns as indicators of conflict in controversial topics (Flöck, Laniado, Stadthaus, & Acosta, 2015; Rad & Barbosa, 2012; Suh, Chi, Pendleton, & Kittur, 2007). Regarding wiki talk pages, social network analyses of editor inter-relatedness have been performed, mainly focused on analysing discussants as editors in the article creation process (de Laat,

Lally, Lipponen, & Simons, 2007; Feldstein, 2011; Matschke, Moskaliuk, & Kimmerle, 2013; Nurmela, Lehtinen, & Palonen, 2000). Thus, there still seems to be an evident lack of empirical research on processes that occur in the background of the article creation, namely within the talk pages as a discussion platform. Therefore, this work's focus is set on analysing processes and the deployment of interventions directly aiming at the talk pages of wiki articles.

1.2 Processes of (collaborative) writing

In general, the composition of a text is not necessarily a straightforward structured task. The required processes differ depending on the goals, the settings, and the individual writer as learner and thus, do not follow an invariant order of production stages. In reference to the Cognitive Process Theory of Writing (Flower & Hayes, 1981), writing is defined as a distinctive set of thinking processes. These are organised hierarchically and can be embedded within any other process (cf. Figure 1.1). As an example, the goal-setting of a writing task can be redefined multiple times during all phases of reading, generating, or editing texts and is not exclusively bound to any specific activity. To successfully achieve one's individual writing goals, the writer should activate monitoring processes such as planning what to write, translating one's thoughts into written text and reviewing text segments. Situational and personal conditions influence these three sub-processes of monitoring, such as the writing task itself, the physical writing environment, individual skills, and existing prior knowledge related to the writing task.

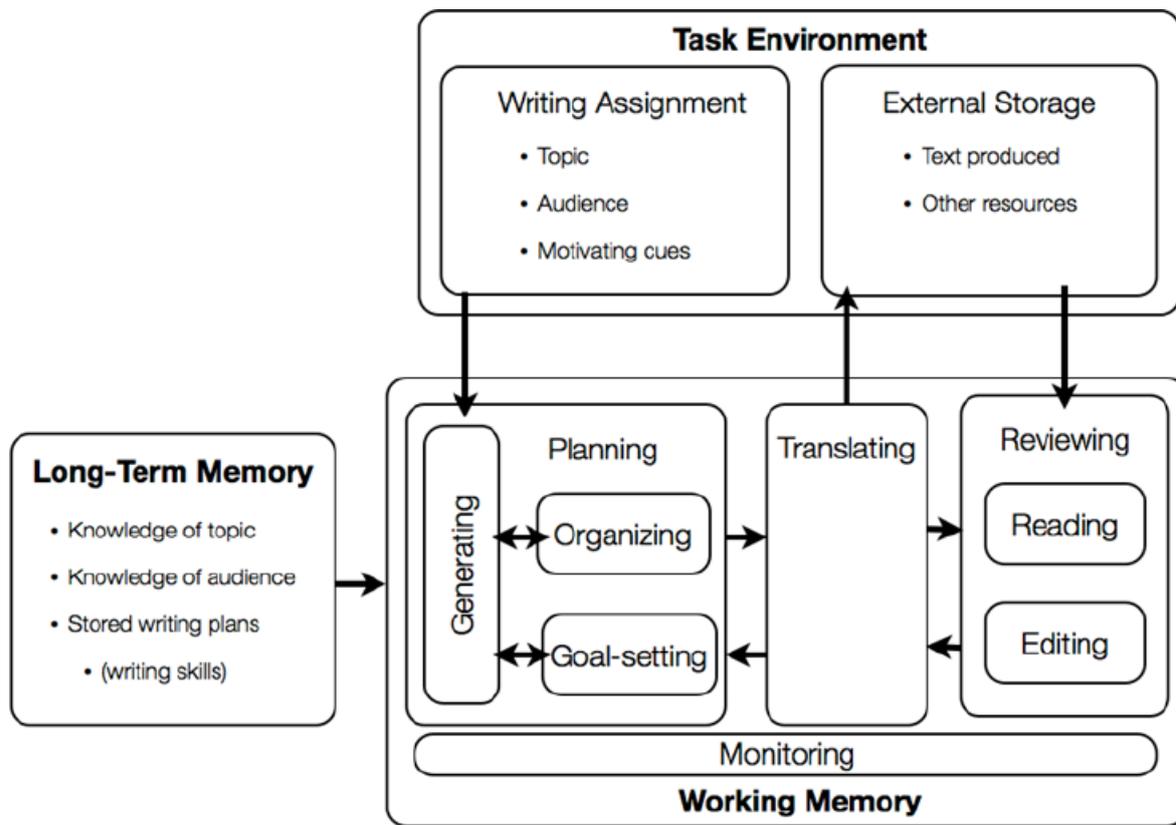


Figure 1.1. Cognitive process model of writing. Adapted from Flower and Hayes (1981)

An individual's writing expertise is closely related to the challenges one might face in a collaborative writing scenario. Wikis seem to be an ideal platform for research in this area, because they were specifically designed for collaborative knowledge construction that is theoretically supported by an unlimited number of individuals. Text production and furthermore the revision of a socially shared artefact can be a challenging task, especially when building and constructing new knowledge within a collaborative setting. Beyond individual strategies and skills for effectively organising one's own writing, operating environments for collaborative writing should provide basic requirements. These include the optional set up of individual roles (e.g. author, editor, or proof-reader) and activity spaces to enable monitoring processes on a group level (Posner & Baecker, 1992).

Research in this area has shown that coordination processes in collaborative writing settings especially are crucial for the quality of the resulting shared knowledge artefacts (Erkens, Jaspers, Prangma, & Kanselaar, 2005). To help with some cognitive processes of writing, most wikis provide substantial functionalities for monitoring the writing processes, such as revision support for text segments. A very basic, but even until today essential revision technique for monitoring, has been implemented in most popular wiki software (e.g. MediaWiki). This technique provides direct side-by-side comparisons of two text versions in a revision history. Reviewing text segments is one of the key monitoring processes in a collaborative writing environment. This includes the revision of one's own text segments as well as reviewing the externalisations of others' contributions to the system. For collaborative writing tasks, Scardamalia and Bereiter (1983, 1985) suggested a self-regulated revision model. Authors' main activities are to compare, diagnose and operate (CDO). They apply these activities to a written text on a sentence by sentence basis in a recursive cycle. The application of CDO proposes to compare two text representations: the already written version with a planned revision. This comparison is like the implementations of side-by-side comparisons in revision histories that can be found in most wikis.

Several studies showed that the usage of CDO resulted in more revisions than intended and that higher quality text passages have been produced (De La Paz, Swanson, & Graham, 1998; Graham, 1997) (Scardamalia & Bereiter, 1983, 1986). The overall quality of the whole revised texts for CDO users was only higher in a few studies (De La Paz et al., 1998). Other studies did not find measurable differences in any direction for the resulting overall text quality while using CDO (Alamargot & Chanquoy, 2001). Accordingly, CDO is not necessarily considered a distinct family of theories by itself, but rather as an explicitly scripted procedure for text revision

performed by an individual in a collaborative setting. Because of the partly contradictory evidence that research gathered for writing outcomes in collaborative tasks, there seemed to be a further need to analyse measures for individual contributors in collaborative environments that can also be used for learning purposes.

1.3 Knowledge building and wikis

In its origin, knowledge building was defined as the creation of knowledge as a social product (Scardamalia & Bereiter, 1994). This original definition suits perfectly well to the idea of collaborative generation of socially shared artefacts within wikis. A significant amount of research has been done on how individual learning and knowledge construction processes can be backed by computer-supported collaborative environments like online discussion forums, blogs or wikis (Ioannou, Brown, & Artino, 2015; Lai & Ng, 2011; Mayordomo & Onrubia, 2015). Scardamalia and Bereiter (2003) explicitly distinguished processes of knowledge construction from individual learning. In this earlier tradition of learning and instruction with computers, learning was defined as an internal and unobservable individual process. In contrast to that, knowledge construction was clearly delimited and encompassed the creation or modification of public knowledge, but without the aspects of collaborative interactions. Dillenbourg, Järvelä, and Fischer (2009) considered social interactions as the essence of cognitions in computer-supported collaborative learning environments, because the creation or construction of knowledge can be regarded as a social product (Vygotsky, 1986). These interactions have been further used to advocate the restructuring of education in the sense of shared problem-solving (Scardamalia & Bereiter, 1994). Wikis are representatives of a collaborative platform type that offers many potentials for shared problem-solving and thus, they have been widely used

in various educational contexts (Notari et al., 2016). They enable people from all over the world to relatively simple content creation and editing without the need for major programming skills (Chen et al., 2015). Moreover, they allow a theoretically unlimited number of authors and editors to build up a corpus of knowledge in a set of interlinked web pages (Franklin & Harmelen, 2007). Nowadays, it is easily possible to measure patterns of mutual influences between individuals and socio-technical systems for knowledge construction.

1.4 The co-evolution of knowledge

Cress and Kimmerle (2008) extended the original ideas of knowledge construction (or knowledge building as sometimes used interchangeably) with a social component, grounded in a combination of Piaget's theory of equilibration (Piaget, 1977a, 1977b) and the systemic approach by Luhmann (1984, 1997) where society is described as a comprehensive social system that includes all other social systems. The focus is on a cognitive and social system as well as their mutual influence. The resulting theory of co-evolution describes and explains the interacting processes of an individual's cognitive system (e.g. author or editor) with a socio-technical system (e.g. wikis), both interacting with each other by processes of internalisation and externalisation of knowledge into one or the other system through processes of accommodation and assimilation. These equilibration processes are fundamental to the co-construction of socially shared artefacts which are prevalent in socio-technical systems such as wikis. Piaget (1977b) identified in his constructivist approach that a disturbance of the cognitive balance is a prerequisite for the construction of new knowledge.

Concretely speaking, in the context of knowledge construction with wikis as computer-supported collaborative learning environments, an imbalance can be resolved by internalising information from the social system into one's own knowledge or by externalising knowledge from one's personal expertise into the wiki. Another possibility of resolving an imbalance would be the restructuring of information in the wiki or restructuring one's individual schemata. These processes match to Piaget's (1977b) assimilation and accommodation processes to restore an equilibrium that was disturbed by a socio-cognitive conflict. Within these processes lay potentials for controversies to arise between two or more collaborating individuals. Different collaborative learning scenarios often set the focus on conflicts arising from the clash of different perspectives or differences in prior knowledge of learners, because in these conflicts lay potentials to create interest and maintain motivation (Isaacs & Clark, 1987; Johnson, Johnson, & Smith, 2000). Some recent research that has been conducted on confirming the theorised co-evolution processes has shown promising results by analysing learning outcomes on individual and collaborative levels in wiki-based environments. There have been positive effects on processes of collaborative knowledge co-evolution and individual learning regarding the occurrences of incongruities between different information sources (Kimmerle, Moskaliuk, & Cress, 2011; Moskaliuk, Kimmerle, & Cress, 2009). Beyond that, there were positive effects between internal and external accommodation and assimilation processes (Kump, Moskaliuk, Dennerlein, & Ley, 2013; Moskaliuk et al., 2009), and also confirmatory results on the polarity of information and redundancy of prior knowledge (Moskaliuk, Kimmerle, & Cress, 2012). These studies give supportive evidence for the co-evolution of knowledge processes in wiki-based learning.

To be more precise, they have specifically addressed the construction of knowledge artefacts on the article level of wikis and not explicitly on the level of talk page discussions where users directly share their arguments with each other on a subject matter. Based on different opinions and contradictory or complementary knowledge, controversies between individuals can be further used constructively to promote learning outcomes (Johnson & Johnson, 1985; Lowry & Johnson, 1981). They can trigger deliberation processes to reach a temporary consensus and furthermore lead to the acquisition of new knowledge or broadening and restructuring of already existing knowledge. Such deliberation processes can finally lead to socio-cognitive conflicts that induce a reorganisation and restructuring of cognitions and ultimately manifests in a learning growth for the individual (Bell, Grossen, & Perret-Clermont, 1985; Bell, 2004). In most of the earlier research settings, the emergence of socio-cognitive conflicts between the individual's internal prior knowledge and the externalised knowledge in the wiki played an important key role (Kimmerle et al., 2011; Kump et al., 2013; Matschke et al., 2013; Moskaliuk et al., 2009; Moskaliuk et al., 2012; Papadopoulos et al., 2013; Wichmann & Rummel, 2013; Zheng et al., 2015). But research so far has not systematically provided support for individuals when they experience these conflicts. So far, research has neglected processes of individual information-seeking and selection behaviour on the layer of wiki discussion pages and their potentials for arising conflicts. In the following section, a closer examination on the significance of socio-cognitive conflicts and their relevance for learning and knowledge construction will be provided.

1.5 Socio-cognitive conflicts

Conflicts between existing knowledge in cognitive systems of individuals and social systems like wikis do not need to have a negative impact on a learner. Quite the contrary, it is often the case that the confrontation with opposing information and further elaboration can lead to conducive learning processes and desirable outcomes. Such conflicts emerge when a person's cognitive schemes contradict either another perspective or knowledge base and as a consequence lead to reorganisation and restructuring of cognitive processes, if consensus building is requested or required (Bell et al., 1985). These socio-cognitive conflicts are also of particular significance for the field of computer-supported collaborative learning because collaborating in a group can lead to higher cognitive achievements compared to an individual working alone (Doise, Mugny, & Perret-Clermont, 1975). Conflicting knowledge bases of group members can act as a motivational driving force for equilibration processes in wiki-based environments or similar collaborative writing settings. It is most beneficial and meaningful for learners when socio-cognitive conflicts emerge out of content-related discussions that are mainly led by evidence and based on contradictory information or controversial positions. Learning benefits on the individual and group level can be evoked by encouraging learners to constructively discuss controversies (Lowry & Johnson, 1981). Furthermore, controversial opinions and points of view have the potential to direct the focus of attention and foster interests (Johnson, Johnson, & Tjosvold, 2000) and can lead to longer-term learning successes by forming and presenting newly acquired lines of argumentation (Doise et al., 1975; Mugny & Doise, 1978). Even if the learner is provided with incorrect information that causes a socio-cognitive conflict, restructuring and reassessing processes can still be triggered and thus lead to the attainment of higher cognitive levels (Bell et al., 1985).

Promoting the opportunities of taking another perspective to discussants while contributing to meaningful discussions can foster elaboration processes and trigger situational Epistemic Curiosity (Weinberger, 2003; Weinberger, Ertl, Fischer, & Mandl, 2005). Beyond content-related controversies that are mainly covered in this thesis, predominantly structural or socio-emotional controversies can also be identified in social systems like on-line discussion forums or blogs (Bodemer, Gaiser, & Hesse, 2011; Janssen & Bodemer, 2013). On existing wiki talk pages, a bandwidth of different conflict types can be found, ranging from socio-emotionally driven disputes to significant evidence-led discussions which comprise hidden potential for knowledge construction processes. Highlighting the latter kind of controversies in wiki's underlying discussion threads might guide interested individuals towards essential learning processes based on socio-cognitive conflicts. It is important to note that socio-cognitive conflicts by means of the co-evolution of knowledge model do not inevitably require that individuals must be involved in constant interaction with each other (Cress & Kimmerle, 2008; Heimbuch & Bodemer, 2017). Even simple interactions of an individual's cognitive system with pre-existing contents in a social system that others have generated in a socially shared manner are socio-cognitive in this model (Figure 1.2). This kind of socio-cognitive conflict becomes especially clear in asynchronous systems such as wikis where no contributor has a guarantee to receive direct or indirect feedback by others within a narrow time frame or even at all. Due to the large information mass that can be present on established wiki talk pages, it is also evident that users can easily be overwhelmed and might be unable to assess a source's quality that is involved in a controversy or causing a socio-cognitive conflict. Consequently, wiki users might feel the need for help to find out what a reliable information source could be.

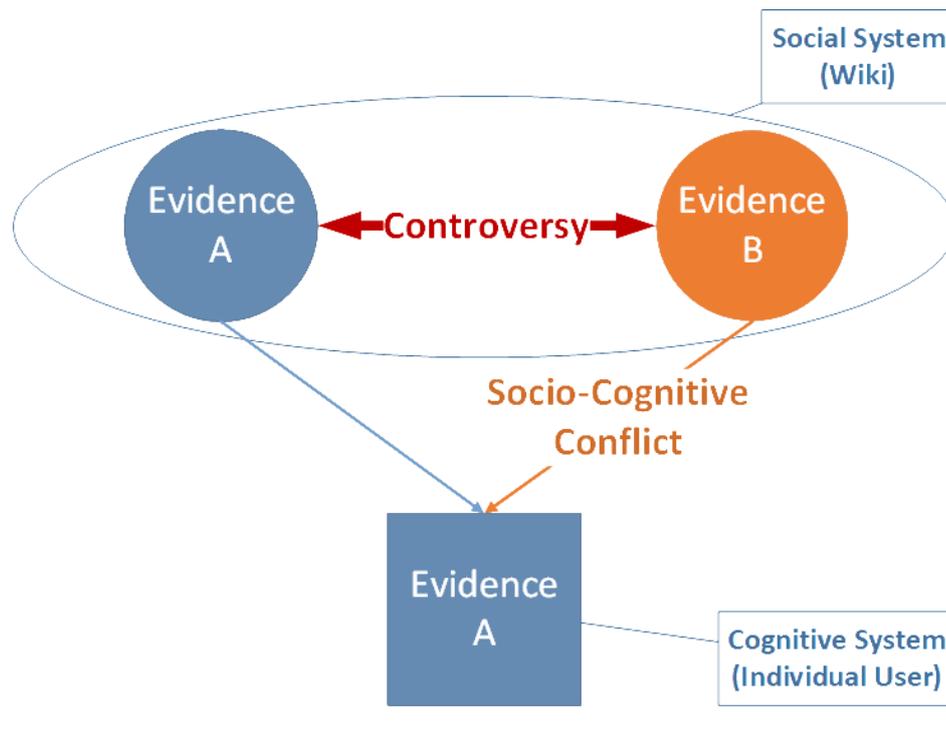


Figure 1.2. Schematic representation of the occurrence of a socio-cognitive conflict, in line with the co-evolution model. An individual with prior knowledge about Evidence A perceives a controversy between discussants in the social system with arguments for Evidence A and Evidence B. The reception of Evidence B in the social system conflicts with the individual's prior knowledge about Evidence A.

1.6 Need for help in wiki-based learning

Seeking help in Web 2.0 environments like Wikipedia is not exclusively bound to a social level; it can also take place on the level of information. A potentially helpful source in this case can be either the information itself or a contributor who provides information (Makara & Karabenick, 2013; Sundar, 2008). When a user assesses information within talk page contributions, they might also assess the information provider, namely the author. Makara and Karabenick (2013) presented a model of a help-seeking process within technology-enhanced learning environments that make use of forums, blogs, or wikis. The framework encompasses seven steps in this process (cf. Figure 1.3), which in full detail are beyond the scope of this thesis.

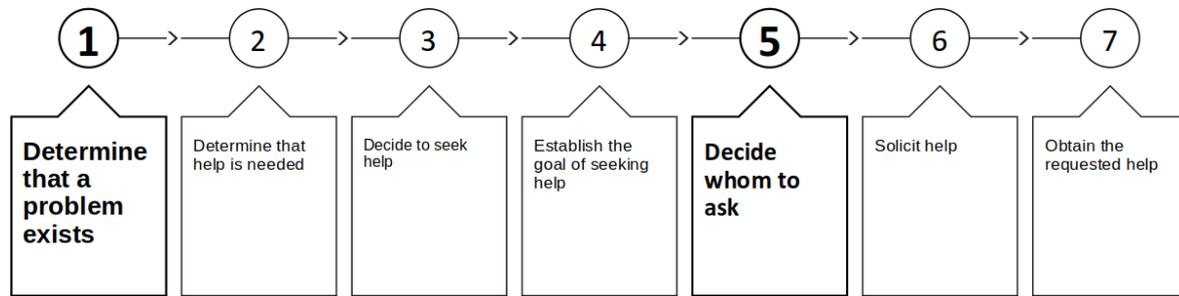


Figure 1.3. Framework of the help-seeking process, adapted from Makara and Karabenick (2013). The stages highlighted in bold represent starting points of guidance measures for wikis as discussed in Experiments 1 and 2 of this thesis (cf. Chapter 2 and 3, respectively).

Research on wiki-based environments that will be later covered in this thesis (cf. studies described in Chapter 2 and 3) can be linked to this framework because they addressed the first two initial steps in the help-seeking framework, i.e. 1) Determine that a problem exists and 2) Determine that help is needed, by implementing representations that aid learners by focusing them on relevant discussions that contain various content-related controversies (Heimbuch & Bodemer, 2017). In another study, subsequent steps in the help-seeking process were addressed regarding the identification of a reliable information source: 5) Decide whom to ask and 6) Solicit help (cf. Chapter 5). If a user of a wiki has already determined that help is needed (“I want to know more about XY, but I’m feeling lost in all these discussions.”) and decided that help shall be sought, one of the subsequent steps in the help-seeking process is a source evaluation of whom to ask and potentially approach for soliciting help. ‘Whom’ can relate to either the provided information or the provider of an information. According to the help-seeking model of Makara and Karabenick (2013), users of technology-supported sources often use expectancy-value heuristics for assessing source quality. A user reviewing a wiki discussion could derive the value of an author through an assessment of the quality of the author’s individual discussion contributions.

In this assessment process, this user might have expectations of a source (author) about how receptive the author would be to requests for help or about the author's community-standing as a discussant. community-standing can be regarded as an indirect measure of the availability of a possible expert in a specific domain, following the idea that an author who is highly involved in the community is more likely to reply to individual on-demand requests (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003; Chang & Chuang, 2011). Users as wiki group members seek information on what is known by others for developing awareness of who knows what (Noroozi, Biemans, Weinberger, Mulder, & Chizari, 2013). To support the establishment of such awareness and share the cognitive labour of for example working on an article, trust and credibility in what is externalised by others becomes important (Lewis, 2003). During this process, it is necessary to assess whose replies and contributions are credible and thus valuable for the user and maybe for the rest of the wiki group as well. Consistently, the conceptualisation of a source's credibility encompasses at least the two key dimensions of trustworthiness and perceived expertise (Berlo, Lemert, & Mertz, 1969; Giffin, 1967; Hovland, Janis, & Kelley, 1953; McCroskey & Teven, 1999; Whitehead, 1968). It has been argued that the execution of heuristics, such as the aforementioned expectancy-value heuristic, is preceded and cued by a subset of affordances which are also relevant for Web 2.0 environments (Makara & Karabenick, 2013). Sundar (2008) has identified several affordances which he subsumed as modality, agency, interactivity and navigability (MAIN). The assignment of a specific modality can cue novelty heuristics for visual representations of information via the perceptual bandwidth, which can be triggered for example through textual, visual, aural or audiovisual modalities (Sundar, 2008). In web contexts, the predominant modalities are text and pictures that are each assigned various levels of credibility, because they vary in degrees of their potentials for noise and deception.

Computer-based media sometimes challenge their users with uncertainty about who or what the true source of information really is. Thus, the agency affordance comes into play as it is connected to credibility considerations around the information source, which can be the content itself, the author as the content creator or the organisation providing the environment. Assigning agency can cue heuristics like bandwagon as a form of social recommendation via measures of collaborative filtering (Buder, Schwind, Rudat, & Bodemer, 2015; Metzger, Flanagin, & Medders, 2010; Winter & Krämer, 2014). In this MAIN model, each component is cueing a number of different additional heuristics via distinct mechanisms that enable quality evaluations before meaningful judgements of a source's value can be assessed. Since one of the research foci of this PhD thesis about wikis lays on the talk page discussions and on the authors of individual discussion contributions, on specific focus was set on the affordances of modality, which is tied to the structure of the medium and agency, which makes the assignment of source in digital media possible. Although interactivity and navigability affordances can be considered as equally relevant, they are beyond the scope of this thesis, which explicitly addressed modality and agency on wiki talk pages to support the assessment of information source quality regarding author credibility and expertise.

1.6.1 Helping to judge an information source

Wikis enable users to perform very influential and drastic changes to the whole environment and its community-generated shared artefacts (Kimmerle, Moskaliuk, Oeberst, & Cress, 2015). Controversial discussions can be used as an extended or supplemental knowledge base to the original article they are related to, but it can be a Sisyphean task to filter what is a good and relevant contribution for oneself, due to the ever-growing base of discussion threads and replies. Therefore, it seems

crucially important that contributors' expertise be easily recognisable for new users. Similarly, contributors' credibility as seen by the wiki community should be visible. Yet, there are no indications of an author's expertise in a certain domain or general expertise represented in wikis. This can become especially relevant if users perceive controversies that are of potential interest for them and want to assess the arguments of one or all the parties involved in the discussion. The lack of effective guidance probably leads a considerable number of users to the decision of leaving the discussion or even the wiki, because they get frustrated with the environment and therefore do not learn or contribute anything new (Capdeferro & Romero, 2012). Wikis are social media platforms with a distinct focus on creating and preserving knowledge, but there are virtually no social components prevalent except for showing user names in discussions and revisions and the possibility to maintain an individual user page. There are no further affordances included in the system to easily assess individual contributions to socially shared artefacts (Notari et al., 2016). The establishment of mutual trust through credibility based on expertise can lead to externalisation of specialised knowledge (Zheng, 2012). It can further enable users to engage in deeper elaborations on materials and as well challenge the opinions of others (Edmondson, 1999). Such challenges of established schemata can be desirable for constructive controversies to occur that might induce socio-cognitive conflicts (Mugny, Butera, Sanchez-Mazas, & Perez, 1995). It is often useful or even necessary to filter down to what could be most valuable and thus to reduce the complexity of information that computer-based media offer. A valuable addition for not getting lost in the bulky information space of wiki talks is to generate easily understandable visualisations based on collaborative filtering mechanisms (Bobadilla, Serradilla, & Hernando, 2009; Konstan & Riedl, 2003).

Research has shown that media users prefer low level strategies of finding “best fits” over high level strategies for the assessment of source credibility, meaning they often prefer quick and easy solutions over more complex processes of weighing and comparing (Cheng, Liang, & Tsai, 2013; Flanagin & Metzger, 2000; Lucassen & Schraagen, 2012). In online media sources, another low level and very effective filtering strategy is the implementation of social recommendation representations (e.g., star ratings for product quality) (Winter & Krämer, 2014). While blogs represent a form of user-generated and participatory websites, they largely differ from wikis in the kind of content creation. In a blog, a single author writes an article that others can comment on, while in wikis all content like articles and related discussions is collaboratively created. An addition of visual support to increase the salience of users who are experts in a controversial discussion could help the processing of one’s own socio-cognitive conflicts and furthermore support the resolution of a continuing controversy in the form of a consensus. Such supplemental information could be very valuable for new and inexperienced users since most wiki contributors are often not experts in the field they are reading and writing about (Notari et al., 2016). For user-generated content sites like blogs, research suggests that social recommendations can predict which source would be read first and followed further. Users’ subjective assessments suggested that social recommendation cues did not guide their source selection preferences (Winter, Metzger, & Flanagin, 2016). In contrast to that, objective log data showed that highly recommended sources are selected earlier, more frequently and read more intensively. These findings consistently suggest that media users are not willing to invest much time and resources for the validity of an information or source in terms of credibility, trustworthiness, and expertise.

Over-reliance in trust and credibility without effective use of expertise information can be counterproductive for members of a community (Zheng, 2012). This is possible to occur when users are not able to accurately understand the level of expertise of an individual contributor. In other related research on wiki talk page discussions, the focus lay on analyses of direct and indirect effects of minimal obtrusive representations on learning outcomes and on the quality of wiki contributions. This research showed promising positive results of representations tacitly guiding users towards relevant discussions on a subject matter (Heimbuch & Bodemer, 2017; Trocky & Buckley, 2016). Building on this line of research, interest arose in exploring further variables that are directly related to processes of selecting an appropriate source and thus are related to further processes of knowledge construction in wikis. Specifically, the influencing potentials of the Need for Cognitive Closure (Webster & Kruglanski, 1994) seemed to be an obvious choice for further research. This construct is closely related to information seeking and its processing and above all how individuals deal with ambiguities, which are prevalent in typical controversial discussions. People who score high on the Need for Cognitive Closure spectrum tend to base their decisions on heuristics, while low Need for Cognitive Closure individuals prefer more information in situations of uncertainty (Dreu, Koole, & Oldersma, 1999; Schlink, 2009). The relevance of such heuristics becomes especially important because of the previously discussed processes in the help-seeking framework and MAIN model which both describe the prevalence and reliance of certain heuristics when assessing the credibility of information sources. The immediate effects of additional information about authors' expertise and credibility are described in an experimental study that will be presented in Chapter 3. For this experimental study, visualisations of domain-specific expertise were created as a ranking number embedded in a badge and community-standing of a contributor was visualised as

an approval count. There was specific interest in several aspects of how wiki users perceive and subjectively rate presented controversial discussions and the added information itself. Both types of supplemental information use the text and picture modalities that are predominant on web sites, and they were aimed at assigning agency to the community of wiki users with the potential of cueing heuristics that are likely to be processed otherwise depending on individual differences.

1.6.2 Influences of individual differences

Besides the individual levels of experience with a subject matter it has been discussed that individual prerequisites and differences in personal and cognitive variables should be considered in supporting and analysing learning processes (Heimbuch & Bodemer, 2017, 2015b). This subsection focuses on individual key variables that can be very relevant for collaborative learning processes. Because of the versatility of an individual's preferences to learning and knowledge construction, it can also be expected that specific key cognitive variables influence how learners handle controversies in discussions grounded on opposing evidences. The first construct of interest in this regard has already been mentioned in previous sections, (1) namely the Need for Cognitive Closure. This construct and its meaning for the research presented in this thesis will be further elaborated. Furthermore, three other individual variables of relevance will be discussed in the following: (2) Epistemic Curiosity which is also closely tied to the Need for Cognitive Closure, (3) intrinsic motivation, and (4) the prevalence and use of metacognitive strategies. They all are closely related to the individual processing of new knowledge artefacts and dealing with ambiguous situations in learning scenarios.

Need for Cognitive Closure. While socio-cognitive conflicts prompt equilibration processes in accordance with the co-evolution model, there is little to no indication of a direction or the intensity of knowledge construction or other cognitive advancements. In terms of the theory of lay epistemology, knowledge emerges by generating hypotheses that are tested deductively (Kruglanski & Freund, 1983). Under specific circumstances people tend to freeze (i.e. epistemic freezing) their hypotheses and do not participate in further testing if they cannot generate plausible alternative hypotheses or if they do not find any evidence to the contrary. Transferred to processes of knowledge construction and learning, the effort a learner is willing to invest in searching for solutions to a problem during a learning process has been indicated to be influenced by the individual Need for Cognitive Closure. This cognitive variable represents a motivational continuum between the need to acquire a clear answer in an ambiguous situation and the avoidance of simple and unambiguous answers. Various empirical results and discussions of the construct illustrate that the Need for Cognitive Closure can be regarded as a relevant construct in knowledge creation processes (Webster & Kruglanski, 1994; Heimbuch & Bodemer, 2017). The motivation to generate new alternative hypotheses depends on the need for structure, fear of invalidity and need for specific conclusions (Kruglanski & Freund, 1983; Kruglanski & Mayseless, 1987). These motivational determinants have been summarised under the overarching construct Need for Cognitive Closure that can either be measured as a personality variable or induced in a specific situation (Kruglanski & Webster, 1996; Webster & Kruglanski, 1994). The individual's Need for Cognitive Closure is best described as an epistemic motivation with regard to the processing and judging information or as a striving for definite and quick answers to a problem and the avoidance of ambiguity (Kruglanski & Webster, 1996).

The situational Need for Cognitive Closure rises under circumstances when predictability or immediate action play an important role and when deeper processing and elaboration of information is required and thus associated with an increase in cognitive load (Kruglanski & Webster, 1996; Webster & Kruglanski, 1994). The latter can be the case if a person experiences a task as monotonous and fatiguing. Such mental fatigue (Webster, Richter, & Kruglanski, 1996) favours a higher Need for Cognitive Closure induced by the situation. These effects get weaker if a person must justify the results and therefore has a higher fear of invalidity. A high dispositional Need for Cognitive Closure is closely related to a reduction in the internal hypotheses building process (Kruglanski & Mayselless, 1987; Kruglanski & Webster, 1996). Persons with a high Need for Cognitive Closure experience ambiguity as unpleasant and tend to be more reluctant in accepting additional information that is inconsistent to their current state of knowledge. These persons are generally more susceptible to cognitive biases such as primacy effects, stereotypes and anchoring heuristics (Dijksterhuis, van Knippenberg, Kruglanski, & Schaper, 1996; Dreu et al., 1999; Ford & Kruglanski, 1995; Kruglanski & Webster, 1996), generate fewer hypotheses and come to final judgements on the basis of easily accessible information and pre-existing knowledge constructs (Ford & Kruglanski, 1995; Webster et al., 1996). These individuals prefer a definite answer in a judgement situation (Schlink, 2009). They are more likely to experience the need for reaching cognitive closure as quickly as possible. Furthermore, it is more relevant to them to maintain a state of closure as long as possible. Conversely, a low Need for Cognitive Closure leads to more information seeking in situations of uncertainty and appreciation of ambiguity (Schlink, 2009). In wiki-based learning, low Need for Cognitive Closure individuals are more likely to seek purposefully for additional in-depth information about a topic in an ambiguous situation. Although this construct is closely tied

to inter-individual differences in learning and knowledge construction, it has yet rarely been covered in studies technology-enhanced or computer-supported collaborative learning. Therefore, this gap is explicitly addressed in this thesis with two experimental studies on learning and knowledge construction in wiki-based environments supplemented by two distinct kinds of collaboration scripts (cf. Chapter 4) and another experimental study building upon these studies where the individual Need for Cognitive Closure plays a key role (cf. Chapter 5).

Epistemic Curiosity. The individual Need for Cognitive Closure can be helpful to obtain a deeper understanding of how ambiguity is processed during the potential emergence socio-cognitive conflicts. Beyond that, the individual manifestation of Epistemic Curiosity may predict information search patterns to further support even deeper evaluation processes of such conflicts (Berlyne, 1954a). Epistemic Curiosity by itself is not coercively a trait but more likely an actual state. According to Berlyne (1954b, 1954a), curiosity by itself is a motivational state and is always specific to finding answers to emerging questions. In these origins, Epistemic Curiosity essentially encompasses the two dimensions of diversive exploration and specific curiosity. Specific Epistemic Curiosity is directed towards a defined task or problem solution that is triggered by a question and can be satisfied by seeking new information or knowledge artefacts that provide a meaningful answer. In collaborative situations comprising controversies that may induce individual socio-cognitive conflicts, the level of specific Epistemic Curiosity could predict what kind of information a learner would seek to find the best possible solution to a concrete problem. Persons with a high situational level of Epistemic Curiosity tend to place more importance on the perceived ease of use and enjoyment of a socio-technical system while searching for new knowledge to solve a problem (Koo & Choi, 2010).

As a result, Epistemic Curiosity correlates highly positive with a smaller number of perceived knowledge gaps in individual learners (Litman, Hutchins, & Russon, 2005). In contrast, diversive exploration arises if a person experiences a lack of stimulation or boredom and thus starts seeking any kind of new stimulus material (Litman & Spielberger, 2003). This rather exploratory kind of curiosity arises out of an information deprivation state and was later termed as D-type curiosity by Litman (2008) who expanded on the original theories and developed several measurements of curiosity. Epistemic Curiosity as a construct can be seen as a complement to the Need for Cognitive Closure where a person with a high need has more likely the desire for a rather unspecific answer in order to reduce confusion and ambiguity (Kruglanski, 1990). Both personality variables can be regarded as intentional states that trigger cognitive processes. Furthermore, they can initiate actions that can be relevant for learning in general and more specific when persons are collaboratively participating in knowledge construction with wikis, where controversies and cognitive conflicts are very likely to occur due to its specific structures.

Intrinsic Motivation. The attention of learners can be directed in such a way that it focuses on the stimuli that promote learning, for example by arousing the curiosity in them to discover new knowledge (Berlyne, 1954a) and to actively participate in shaping it (Scardamalia und Bereiter, 2006). The relevance of a learning task is determined by the subject matter as perceived by learners. It should be promoted by concentrating on learning processes rather than results, so that learners are given the opportunity to learn aspects of knowledge interesting for future tasks together with others and to take responsibility for their own goals and those of others (Keller, 1987a, 1987b).

Accordingly, certainty of success contributes to perseverance and learning success and can be promoted by teachers by removing the fear of possible failure from the learner. The learning requirements should be clearly presented by the instructor and realistically expected by the learner so that the learner feels that he or she has control over the learning progress and success. Satisfaction with learning outcomes should be encouraged by stimulating intrinsic rewards through positive feedback and encouraging the use of newly acquired skills, so that learners do not feel they are learning and being rewarded for extrinsic, uncontrollable reasons, but feel joy and pride when they have solved a problem (von Glasersfeld, 1992). To make use of the potentials collaborative learning offers, collaboration should meet some individual requirements of learners, including motivational processes. Effective learning depends on intrinsic motivation. According to the theory of self-determination by Deci (1992), this kind of motivation is high when collaborative tasks meet the optimal challenges. Furthermore, positive feedback can increase the feeling of competence and self-determination and thus also the intrinsic motivation. Rewards and competitive thinking can also motivate behaviour, but this motivation is called extrinsic. Within the theory's framework, the need for competence and self-determination were identified as fundamental prerequisites for intrinsic motivation. Perceived competence, social affiliation and autonomy in the execution of actions can lead to the feeling of self-confidence and effectiveness. Consequently, behaviour is likely to be internalised and thus becoming a part of the person. Too much control inhibits the perceived sense of autonomy and thus also the motivation (Deci & Ryan, 1985; Deci & Ryan, 2003). Working with wikis requires great initiative and individual activity and therefore a high level of self-motivation.

In recent research on learning and knowledge construction, motivational differences were discussed as interesting latent construct for different learning scenarios. Results were insofar regardless of the settings whether research was conducted in face-to-face classrooms or computer-supported collaborative learning situations (Leppink, 2010; Leppink, van Gog, Paas, & Sweller, 2015). Learner actions that are motivated intrinsically can be a key to effective learning (Deci & Ryan, 1985; Nakamura & Csikszentmihalyi, 2014). Intrinsically motivated persons frequently feel the need to be socially connected to a community while maintaining high levels of personal autonomy and self-determination (Deci, 1992). This does not necessarily mean that intrinsic motivation can only be experienced in isolation, it can also be triggered by external sources in a formal or informal learning setting (Wilde et al., 2009). In principle, wikis offer their users open spaces and sufficient leeway regarding self-chosen thematic foci for intrinsically motivated learning. This in turn requires consistent willingness to show individual initiative which can be initially triggered by the environment itself (Dillenbourg et al., 2009). The opportunities to introduce own prior knowledge and experiences into such a social system can increase an individual learner's motivation. Highly intrinsic motivated learners can be central actors in meaningful socio-cognitive discourse that can be beneficial for more elaborated learning (Rienties, Tempelaar, Van den Bossche, Gijsselaers, & Segers, 2009). Motivation in group learning activities is an essential component that is still not well enough covered in research, equally whether analysing individual or collaborative learning processes (Järvelä, Volet, & Hana Järvenoja, 2010). Therefore, in wiki-based learning where students have different degrees of freedom depending on the task, motivational states can be a relevant influencing factor of successful learning.

Metacognitive strategies. Closely related to motivational states is the concept of metacognition which can be defined as reflective and regulative knowledge about the own knowledge (Efklides, 2008; Flavell, 1979; Kuhn, 2000). Metacognition is often regarded as an individual and conscious process that can also unconsciously operate on a social level during co-regulation in collaborative settings (Efklides, 2008). It is connected to first-order cognitions by monitoring and control functions. These metacognitive functions are helpful for self-regulated learning processes, where self-motivation and the choice of appropriate strategies become relevant (Zimmerman, 2002). Essentially, two distinct levels of metacognitive processes are considered: the object- and meta-level. Both levels interact and exchange information via monitoring and control functions. The meta-level is informed by monitoring processes with information about cognitive activities on the object-level (cf. Figure 1.4). Control functions ensure that meta-level orders reach the object-level (Veenman, Van Hout-Wolters, & Afflerbach, 2006). Besides these levels, metacognition encompasses three involved facets: metacognitive knowledge, metacognitive experiences, and metacognitive skills. Metacognitive knowledge includes explanatory knowledge about tasks, strategies and goals of others or oneself (Flavell, 1974).

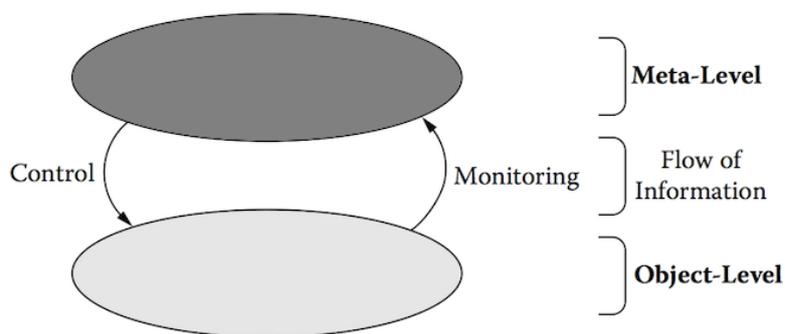


Figure 1.4. In this model representation, metacognitions (meta-level) and cognitions (object-level) are interrelated with each other, sharing information with each other by means of monitoring and control processes. Figure adapted from Nelson and Narens (1990).

Metacognitive experiences contain everything that a person knows and feels during the execution of a task (Efklides, 2006). Experience makes the state of cognitions conscious and trigger control processes that serve to the goals of self-regulation. Both metacognitive knowledge and metacognitive experiences are associated with monitoring functions. In contrast, metacognitive skills represent knowledge of practical value that is triggered by the other two facets of metacognition. The use of metacognitive strategies can be crucial for successful learning, although their effectiveness is not absolutely certain (Berthold, Nückles, & Renkl, 2007). Metacognitive processes such as planning, monitoring and evaluating learning processes is often regarded as time-consuming and painstaking. As a possible consequence, existent skills are not used effectively despite the availability of metacognitive strategies (Neuenhaus, Nora, 2011). In wikis, where by default at the very most rather loose guidance measures for any of these aforementioned metacognitive processes are implemented, learners with less distinct metacognitive strategies might experience collaboration with wikis as unpleasant and could create reluctance to participate (Capdeferro & Romero, 2012). Thus, all the previously portrayed constructs will be examined as possible influences and determinants to identify the degree of additional support a learner needs as potential guidance while observing or being involved in socio-cognitive conflicts during knowledge construction and learning within wiki environments.

1.7 Guiding and supporting learners

Using asynchronous collaborative socio-technical environments like wikis for learning purposes can be a challenging task for its users. Providing learners with media and letting them freely collaborate does not automatically promote systematic learning processes and is dependent on an interplay of numerous variables such as the task itself, characteristics of the group and its individual members or the underlying collaboration media (Stahl, 2006). It has been shown that missing objectives and a lack of structure is problematic for productive interactions and outcomes in a collaborative setting (Bromme et al., 2005). Thus, guiding structured learning and communication processes is essential for the effectiveness of computer-supported collaborative learning settings (Fischer, Kollar, Stegmann, & Wecker, 2013).

1.7.1 Implicit guidance

As previously mentioned regarding the missing of objectives during wiki-based learning, several supportive interventions have been designed and evaluated with a primary focus on the writer's task environment. Examples include the explicit setting of a deadline or the definition of specific goals for an individual's writing task. Such interventions led to a set of extensions and design principles to facilitate group writing and constructivist learning processes for students in higher education (Kasemvilas & Olfman, 2009; Zheng et al., 2015). Regarding the issue of missing structure, socio-cognitive conflicts are in general well suited for triggering individual elaboration and collaborative discourse, explicitly by suggesting separate roles and collaboration processes (Kollar, Fischer, & Hesse, 2006; Heimbuch et al., 2014) or implicitly by providing representational and informational guidance through signalling measures (Cashen & Leicht, 1970).

The latter class of support measures are particularly suited to be embedded within informal information and learning environments as they do not interfere with self-regulated learning processes (Beers, Boshuizen, Kirschner, & Gijsselaers, 2005; Bodemer, 2011; Dillenbourg, 2002). Representational guidance (Suthers, 2001, 2003) impacts processes and task performance in a collaborative environment by providing salient cues. It was shown that representational guidance can nudge the discussion of controversial evidence (Suthers, 2001) and can lead to learning outcomes of higher quality (Janssen, Erkens, Kirschner, & Kanselaar, 2010). Cognitive group awareness tools that were built upon these principles were evaluated to be beneficial for learning processes and outcomes in various settings (Bodemer & Dehler, 2011; Janssen & Bodemer, 2013). They guide collaboration behaviours by providing individuals information on other learners' knowledge or opinions. Research on cognitive group awareness tools in online discussions with many controversial topics showed that they can highlight high quality arguments and opposing points of view which both can lead to an improved perception of minority opinions, a higher frequency of conceptual change and better learning outcomes (Buder & Bodemer, 2008; Buder et al., 2015). Visualising controversies on different points of view or contradictory evidence can also support the emergence of socio-cognitive conflicts which then provide the potential of leading to restructuring processes in one's own cognitive system as well as within the socio-technical system which is used for knowledge construction. However, to gain deeper insights of the underlying processes in an individual contributor, it should also be considered that individual differences in variables relate to the processing of information and dealing with conflicts (cf. Section 1.6.2). Wiki talk pages can serve as meaningful sources of additional knowledge that no author or editor has integrated or referenced in an article.

With this thesis, the provision of additional support for externalisation and internalisation processes of conflicting knowledge shall be analysed. It is believed that enhancing an individual's focus on most relevant aspects of a conflict before they revise an existing text can foster the externalisation of more relevant information from the social system into a cognitive system and subsequently result in learning gains through optimised internalisation (cf. Chapter 2). Furthermore, in Chapter 3 a first exploration of general effects of visual credibility information added to wiki talk page discussions will be provided. The general interest was in how additional information about author's expertise and community-standing are perceived by users on several dimensions from a user experience perspective. Both experimental studies as presented in Chapter 2 and 3 were designed to address various stages in the help-seeking framework as previously discussed in Section 1.6 and represented in Figure 1.3.

1.7.2 Explicit guidance

One of the key components to meaningful interactions between participants in collaborative settings is that the interaction is clearly aimed at a specified learning goal or task. Thus, it can be helpful to predefine and provide certain rules or guidelines for specific tasks, such as creating or updating wiki articles and their corresponding discussion pages (Bromme et al., 2005). This can be rooted in the observation that persons working collaboratively on the same document often implicitly take over separate roles, activities and make use of certain strategies that are not necessarily best suited for effective collaboration (Posner & Baecker, 1992). Especially in high-level collaboration, a certain level of guidance is necessary to exhaust the full potentials of collaborative learning (Weinberger, Stegmann, Fischer, & Mandl,

2007). Such guidance can be provided with collaboration scripts where the activities of individuals in a social system are coordinated and optimised. A script is a set of instructions specifying the group formation, modes of interaction and task management between collaborators that can be distinguished by the level of structuring degree (Kollar et al., 2006). Choosing a desired complexity while designing a script is a trade-off between effective structuring and over-scripting (Dillenbourg, 2002), which needs to be considered by the designer of collaboration scripts. Positive evidence could also be found for scripts with a special focus on article editing and revising. This ultimately led to more coherent articles and less inaccurate or incorrect articles (Wichmann & Rummel, 2013). Evidence of this research suggests that additional scripts applied to wikis can improve the overall quality of knowledge artefacts and coordination processes of learners. In other related research, it has also been shown that a certain level of coercion in the implemented collaboration script is required (Papadopoulos et al., 2013). To make use of the potentials that participatory learning environments provide to different learner types, individual prerequisites should be considered when designing learning scenarios that are aimed at collaboration. Very explicit instructions requesting to externalise mental representations of learned content and actively participate in the revision process can be beneficial. This is especially true for novice learners who are less familiar with a domain and thus need more guidance than more experienced learners (Kirschner, Sweller, & Clark, 2006). Large wikis such as Wikipedia have created inherent proposals for how to create and revise articles. One specific instructional set has emerged that can be referred to as a script for collaborative writing in wikis. Wikipedia's so-called "Be Bold, Revert, Discuss" (BRD) cycle promotes a high frequency of individual article edits without the need of immediate coordination when changes are about to be made (Wikipedia, 2017a).

It was mainly incorporated into Wikipedia as a low-barrier workflow for article creation with the aim to engage many individual contributors. But it can also be used as a collaboration script approach for low-level coordination of knowledge construction processes. The first step in this BRD script shall encourage content creators to be bold. It means that they are asked to perform an edit to an existing article or create an entirely new article, if one feels the desire to do so. The second step “revert” encompasses guidelines on how to deal with one’s edited or newly added content being revised by other users. The last step of discussing content revisions suggests that at least three revisions to one’s edits should be accepted before a discussion should be initiated. These steps are not necessarily designed to be followed in an invariant order, although flow chart representations as used in Wikipedia and alike might suggest a specific order of events. The BRD script encourages primarily a high participation rate without the need for discussion. In the research conducted as part of this thesis, it is thought to be quite important to set a focus on discussions since they provide much potential for learners in terms of beneficial socio-cognitive conflicts. Thus, in this thesis an alternative approach is proposed with a different script that is more in line with current research findings of the learning sciences. It was mainly inspired by wiki-based learning experiments that targeted the improvement of writing and more effective coordination (Papadopoulos et al., 2013; Wichmann & Rummel, 2013). With such an alternative approach to Wikipedia’s BRD proposal, it was aimed to engage participants to discuss any planned article edits and revisions upfront before changes to a document will be performed, resulting in a script called “Discuss, Deliberate, Revise” (DDR). For both scripts, BRD and DDR, detailed text instructions were provided in mandatory tutorials to participating learners as part of the experimental studies (Heimbuch et al., 2014).

In Figure 1.5 the core stages of both collaboration scripts are represented. The visual script representations were solely used as permanently visible reminders. The first step in this DDR script suggests editors to take planned article edits first to the corresponding discussion space. In the following second step the deliberation process shall be initiated where the community is sought to find a consensus about the proposed wiki edits. And finally, in the last step after having reached a temporary consensus the article revision can be performed. Because of these two differing script approaches of BRD and DDR, the potentials of these two collaboration scripts that were designed with different objectives in mind were examined in laboratory and field settings (cf. studies presented as Experiments 3 and 4 in Chapter 4).

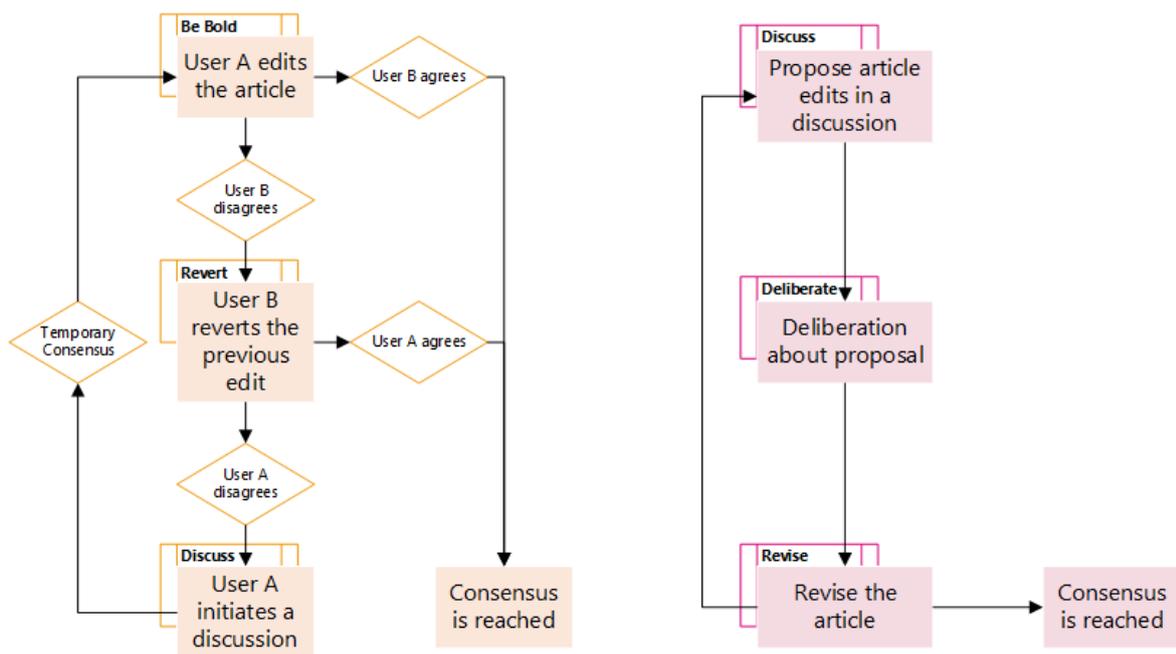


Figure 1.5. Representations of the Be Bold, Revert, Discuss (left) and Discuss, Deliberate, Revise (right) scripts main steps as used in Experiments 3 and 4 (cf. Chapter4) as a flow chart. In Experiment 5 only the Discuss, Deliberate, Revise was used cf. Chapter5.

1.8 General Research Questions

The main research interest within the framework of this thesis was to investigate the potentials of additional scaffolds in terms of visual controversy awareness information and collaboration scripts to influence and guide individual processes and outcomes in wiki-based learning. During the chain of events of externalising and internalising knowledge from or to either system can induce socio-cognitive conflicts between prior knowledge in an individual's cognitive system and the perceived controversial exchange of opposing evidence in the wiki as social system. Specifically, the studies aimed at factors that can facilitate individual learning and knowledge construction within social systems like wikis. All studies had in common that the core independent variable was some variation of additional guidance for wikis. In the following paragraphs, the numbering of the research questions refers to the respective subsequent chapters of this thesis describing the experimental studies. A general model overview of anticipated influences between specific variables and overarching constructs that were analysed in a series of five experimental studies that were conducted for this thesis is summed up after the specific paragraphs in Figure 1.6 at the end of this section.

Controversies and wiki activities. Research on cognitive group awareness tools has shown potentials in different contexts to foster learning processes and enable deeper elaboration (Bodemer, 2011; Buder & Bodemer, 2008; Engelmann, Dehler, Bodemer, & Buder, 2009; Janssen & Bodemer, 2013). Visualisations of information that target at the representation of conflicting knowledge constellations or opinions of particular groups in online discussions can guide learners towards new points of view or to focus on yet unknown facts (Buder & Bodemer, 2008; Buder et al., 2015). Although previous research on cognitive group awareness has discussed making

use of controversies in general, this research did not investigate any visualisations on the occurrence of controversies and the status of resolution yet.

RQ2.1: Thus, specific interest has grown in answering the question: To what extent does visual controversy awareness information influence individually relevant learning-related processes in wiki-based environments?

Controversies and learning outcomes. Particularly because of the partly contradictory evidence that research gathered for writing outcomes in collaborative tasks, there seemed to be a further need to analyse supportive measures for individual contributors in collaborative environments. Possible interventions could be implemented as additional support to already addressed explicit procedures such as CDO (cf. Section 1.7). Since wikis already provide the essential functionalities for text revision, one aim could be to support wiki writers as learners with a preceding step that enables them to focus on the most relevant aspects of their writing task. Cues to what contents are most relevant could allow writers to perform fewer, but more efficient and effective comparisons, more purposeful diagnoses by emphasising the specific kind of problem and thus, finally resulting in a better operation manifested in higher quality texts. Beyond that, when controversies are grounded on evidence rather than opinions socio-cognitive conflicts about contradictory opinions can be induced (Johnson, Johnson, & Tjosvold, 2000; Lowry & Johnson, 1981). The occurrence of such conflicts and its confrontation can trigger equilibration processes that can be beneficial for learners (Bell et al., 1985; Mugny & Doise, 1978). First qualitative content analyses of the present data showed promising direct effects on the quality of individual replies as well as indirect effects on the quality of individually performed article edits (Heimbuch & Bodemer, 2016a).

RQ2.2: Therefore, one question to investigate arose: To what extent do learners use visual controversy awareness information on evidence-led discussion threads to foster quantifiable learning outcomes?

Expertise and Credibility Wikis enable users to perform very influential and drastic changes to the whole environment and its community-generated shared artefacts (Kimmerle et al., 2015). Controversial discussions can be used as an extended or additional knowledge base to the original article they are related to, but it can be a Sisyphean task to filter what is a good and relevant contribution for oneself, due to the ever-growing base of discussion threads and replies. Therefore, it seems crucially important that contributors' expertise be easily recognisable for new users. Similarly, contributors' credibility as seen by the wiki community should be visible. Yet, there are no indications of an author's expertise in a certain domain or general expertise represented in wikis. This can become especially relevant if users perceive controversies that are of potential interest for them and want to assess the arguments of one or all the parties involved in the discussion. The lack of effective guidance probably leads them to the decision of leaving the discussion or even the wiki, because they become frustrated with the environment and therefore do not learn or contribute anything new (Capdeferro & Romero, 2012).

RQ3.1: Does the implementation of additional author information about a discussants' expertise and community-standing aid readers of wiki discussion threads in their individual assessment processes?

For user-generated content sites like blogs, social recommendations have been shown to have some inherent potentials to predict which source of information would be read first and followed further for in-depth investigations. Subjective

assessments of users of such participatory sites suggested that social recommendation cues would not guide their source selection preferences (Winter et al., 2016). In contrast to that, the more objective log data, as recorded by the experimental environments, showed that highly recommended sources are selected earlier, more frequently and read more intensively. These findings consistently suggest that media users are not willing to invest much time and resources for the validity of an information or source in terms of credibility, trustworthiness, and expertise. Overreliance in trust and credibility without effective use of expertise information can be counterproductive for members of a community (Zheng, 2012). This is possible to occur when users are not able to accurately understand the level of expertise of an individual contributor.

RQ3.2: How does additional author information about domain-specific expertise and community-standing influence the information selection behaviour of users in wiki discussions?

Since this experimental study was originally designed as being a first exploration of general effects of novel visual additions to wiki talk page discussions, it was as well of general interest how a new kind of wiki author information that is added to the talk pages is perceived by users on several dimensions from a user experience perspective.

RQ3.4: How do users rate the additional author information implementations about domain-specific expertise and community-standing in terms of their personal user experience?

Influences of individual differences. Socio-cognitive conflicts are possible to occur in wiki-based knowledge construction, both in the individual cognitive system

and in the socio-technical system through processes of externalisation and internalisation (Cress & Kimmerle, 2008). General processes of accommodation and assimilation that are triggered by conflicting knowledge have been identified in both systems (Kimmerle et al., 2011; Kump et al., 2013; Moskaliuk et al., 2009; Moskaliuk et al., 2012). In wiki-related research to date, individual differences regarding the processing of ambiguous information and new perspectives in controversial discussions have not been reported.

RQ2.3: With this study, also the following question should be explored:

To what extent do one's personal Need for Cognitive Closure and Epistemic Curiosity affect knowledge construction processes?

Regarding the assessment of credibility and perception of expertise in wikis, exploring variables that are directly related to processes of selecting an appropriate source seems evident. Specifically, the influencing potentials of the Need for Cognitive Closure should be analysed. This construct is related to information seeking and processing and how individuals deal with ambiguities (Webster & Kruglanski, 1994), which are prevalent in typical controversial discussions. People who score high on the Need for Cognitive Closure scale tend to base their decisions on heuristics, whereas low Need for Cognitive Closure individuals prefer more information in situations of uncertainty (Dreu et al., 1999; Schlink, 2009).

RQ3.3: How is the perception of controversial wiki discussions influenced by the individual Need for Cognitive Closure when additional author information is present or absent?

In a further experimental laboratory study, two different collaborations scripts were to be tested in a wiki-based learning environment – (1) Wikipedia's proposal

Be Bold, Revert, Discuss (BRD), in contrast to (2) an alternative proposal *Discuss, Deliberate, Revise* (DDR). The former script is directed at externalising as much knowledge from an individual's cognitive system into the wiki as social system. The latter script aims at fostering deeper elaboration processes by encouraging discussions prior to the externalisation of knowledge into the wiki and has already shown to be beneficial for mentally integrating different perspectives about a topic. To further support and extend these findings a follow-up field study was conducted as a supplement to a lecture on research methods and statistics. It was expected that collaboratively working on a wiki in an educational context is most beneficial for students when they are first guided into discussions rather than individual encouraged to edit individually without further coordination. In both studies, there was interest in researching potential direct effects of the collaboration scripts. For the field experiment, also data on additional influencing variables were collected that can be related to learning and knowledge construction processes to gain deeper insights about potential indirect conditional effects.

RQ4.1: Does a DDR script have a greater positive impact on learning as well as on the wiki article quality than Wikipedia's proposed BRD script?

While the first research question and hypotheses in this area will be confirmatory, due to a relatively large body of research regarding collaboration scripts and their effects, the second research question is of a more exploratory nature (cf. Chapter 4 for the specific hypotheses that are presented in dedicated sections). The body of research in computer-supported collaborative learning is rather small regarding some variables of interest that were discussed in this thesis, namely the Need for Cognitive Closure, metacognitive strategies, and intrinsic motivation. Consequently, there

was additional interest in exploring the potential effects of individual differences on outcome and process variables.

RQ4.2: To what extent do the individual variables influence the effects on knowledge construction and the frequency of wiki activities in the talk page discussion forums?

Guidance interacting with the Need for Cognitive Closure. Building upon the positive results of previous experimental studies, as will be presented in more detail in Chapters 2 and 4, distinct kinds of implicit and explicit guidance implementations for wiki-based learning environment have already been analysed (Heimbuch & Bodemer, 2017). Both kinds of guidance showed several positive effects on knowledge test scores and wiki contribution quality as well as potentials for more purposeful interactions within the wiki environment. On the one hand, making controversial discussions more salient to wiki users can be achieved by implicitly guiding learners towards those contents. It provides the potentials of receiving conflicts between other wiki participants as well as to induce socio-cognitive conflicts in the individual user itself. On the other hand, a more explicit collaboration script proposal aims at fostering deeper elaboration processes by encouraging discussions prior to the externalisation of knowledge into the Wiki and has already shown to be beneficial for mentally integrating different perspectives about a topic.

RQ5.1: How do (1) implicit guidance with controversy awareness highlights and (2) explicit guidance with a DDR collaboration script affect processes and outcomes in wiki groups?

Previous research has also provided several indications that the Need for Cognitive Closure might have an impact on learning-related processes and outcomes. To

further support and extend on these findings a follow-up laboratory study was conducted to compare the effects of one implicit (controversy awareness status highlights) and one explicit guidance (DDR collaboration script) implementation for wiki environments in interaction with the individual Need for Cognitive Closure of learners.

RQ5.2: How does the individual Need for Cognitive Closure influence processes and outcome variables related to learning in the case of (1) implicit guidance with controversy awareness highlights and (2) explicit guidance with a DDR collaboration script?

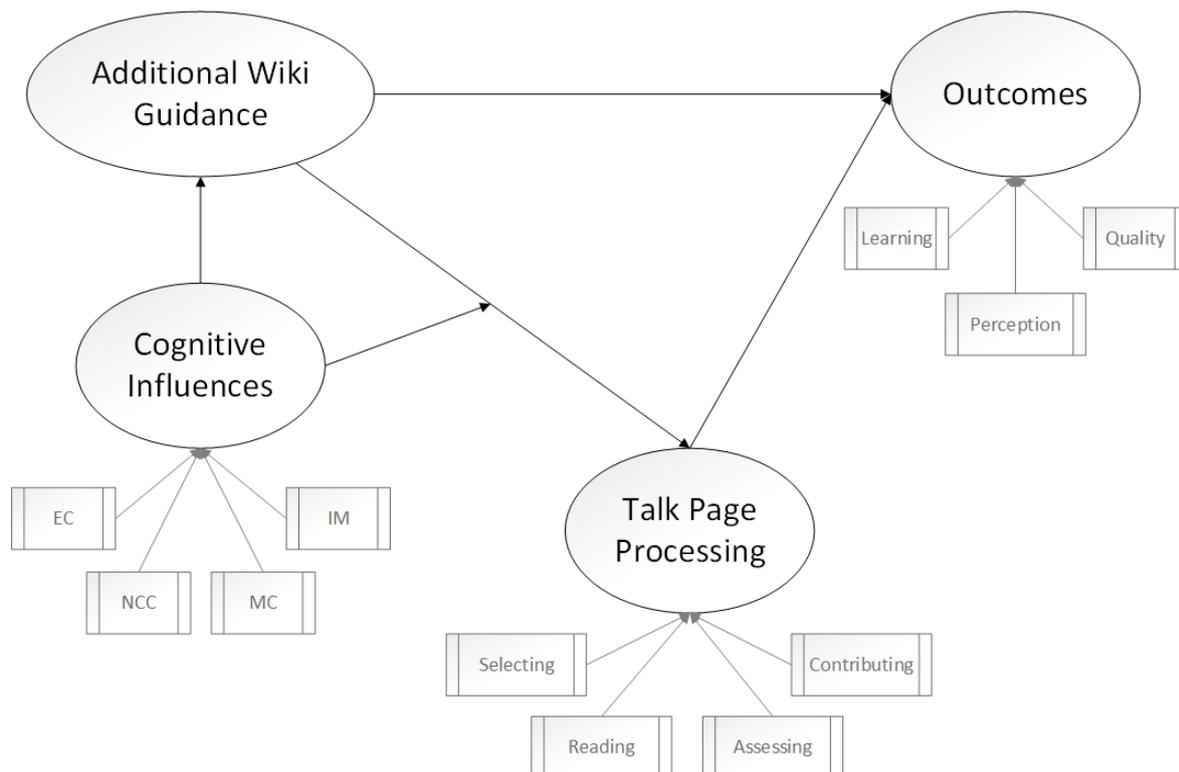


Figure 1.6. In this model hypotheses of the major research questions of the experiments that were conducted for this thesis are represented. Hypotheses encompassed direct and indirect (via mediation and moderation) influences of additional guidance for wikis on process and outcome variables.

1.9 Remarks on Statistical Methods

In the course of realising and completing this PhD project, a range of different statistical methods have been used to analyse the datasets of the five experiments which are described in the following chapters. Some methods are to date not that prevalent in psychological research or represent deviations from defaults as taught in many textbooks about analysing data in Psychology. The relevant methods are briefly described in the following paragraphs. If more details about a statistical method were necessary, due to specific requirements of an experiment, these details are presented in the method or result section of the respective experiment. The datasets and study materials are stored at the *Open Science Framework* and available for download at <https://osf.io/w9v7g/>.

Comparing differences of two or more groups. As a default test for two-group comparisons, Welch's *t*-Tests was applied for the main statistical analyses instead of Student's *t*-Test. Among others, Ruxton (2006) and Delacre, Lakens, and Leys (2017) showed that Welch's *t*-Tests is more robust for smaller and medium-sized samples and consequent an increased likelihood of not meeting either the assumption of normality or equal variances. When more than one assumption of parametric tests was violated (e.g. normality and homoscedasticity of residuals) and/or sample sizes were very small (Experiment 3), non-parametric statistical tests were used for all analyses (cf. Chapter 4).

For group comparisons that involved more than two groups, the Brown-Forsythe test of variances was used as a robust test on the *F*-distribution, if the data violated the homogeneity of variance assumption (Brown & Forsythe, 1974). For analysing multiple dependent variables that were most likely to be correlated by design, multivariate analyses of (co)variance (MAN(C)OVA) were used. In that case, Wilk's λ

was chosen to report as an appropriate test statistic that is adequately powerful to detect statistical group differences when multiple dimensions come to play (Tabachnick & Fidell, 2012).

Analysing conditional (mediation and moderation) effects. Conditional effects of influencing variables on dependent variables (Experiments 1, 2 and 4) (cf. Chapters 2, 3 and 4) were analysed with the *PROCESS* macro (version 2.16.3) for SPSS (version 24.0.0.2) by Hayes (2013). For all mediator and moderator analyses, bias-corrected and accelerated (BCa) bootstrapping with 5,000 samples was applied. Any level of conditional effects reaching statistical significance on $\alpha = 5\%$ are illustrated by Johnson-Neyman plots using the so-called floodlight method. The common pick-a-point approach with simple slopes plots was additionally chosen to illustrate effects that did not meet the criteria of $\alpha = 5\%$, but still are of potential relevance. For all conditional effect analyses, 95% confidence intervals for the effect sizes are reported.

Analysing two-group equivalence. Equivalence hypothesis tests (Experiment 5) were performed with the R package *TOSTER* (version 0.2.5) by Lakens (2017b). The underlying two one-sided *t*-Test (TOST) procedure requires to determine a priori the smallest effect size of interest (SESOI) for the specification of equivalence bounds (Hauck & Anderson, 1984; Schuirmann, 1987). Applying the TOST procedure, in conjunction with a *t*-Test of differences within the null hypothesis significance testing (NHST) framework, can yield four possible outcomes for an effect (cf. Figure 1.7): a) statistically equivalent and not different, b) not equivalent and statistically different, c) statistically equivalent and different, and d) not equivalent and not different.

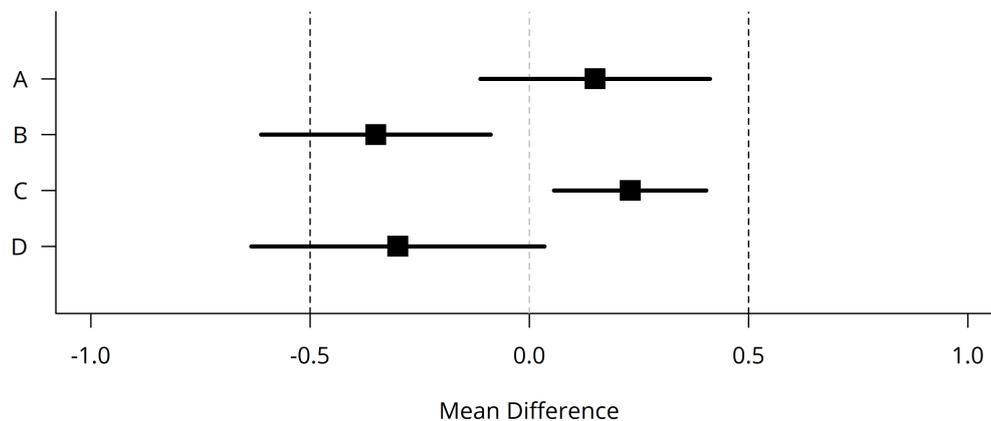


Figure 1.7. Mean differences (black squares) and 90% confidence intervals (horizontal lines) with equivalence bounds $\Delta L = -0.5$ and $\Delta U = 0.5$ for four combinations of test results that are statistically equivalent or not, and statistically different from zero or not. Pattern *A* is statistically equivalent, pattern *B* is statistically different from 0, pattern *C* is practically insignificant, and pattern *D* is inconclusive (neither statistically different from 0 nor equivalent). Figure adapted from Lakens (2017a).

Comparing statistical evidence with Bayes Factors. Bayes Factors were computed with the open source statistics program *JASP* (version 0.8.2.0) developed by The JASP Team (2017) and the R package *BayesFactor* (version 0.9.12-2) by Morey and Rouder (2015). The addition of Bayes Factors is also suggested when smaller to moderate sample sizes are present (Experiments 1, 2, 3 and 4) (cf. Chapters 2, 3 and 4) because approximately 50% smaller samples are required for reliable factors to converge compared to the sample requirements for adequately powered studies within the frequentist framework (Schönbrodt & Wagenmakers, 2017).

For all Bayesian hypothesis tests that involved two-group comparisons an informed prior distribution of $t(0.35, 0.102, 3)$ was used as discussed by Gronau, Ly, and Wagenmakers (2017). For general linear model analyses, a default r scale parameter of $r = .5$ for ANOVAs and $r = .35$ for regression models was used (Morey & Rouder, 2015). When the analyses in the following chapters estimate Bayes Factors

in favour of the alternative model $BF_{10} > 100$, the logarithmic $\log BF_{10}$ will be reported due to better readability. Although Bayes Factors are continuous parameters, a rough heuristic classification aid is provided in Table 1.1.

Table 1.1

A classification scheme for Bayes Factors as suggested by Lee and Wagenmakers (2013).

Bayes Factor	BF_{10}	BF_{01}
> 100	Extreme evidence for H1	Extreme evidence for H0
30 – 100	Very strong evidence for H1	Very strong evidence for H0
10 – 30	Strong evidence for H1	Strong evidence for H0
3 – 10	Moderate evidence for H1	Moderate evidence for H0
1 – 3	Anecdotal evidence for H1	Anecdotal evidence for H0
1	No evidence	No evidence

Reporting of effect sizes. Effect size measures that can be represented by proportions of variance (η^2 , R^2) are reported with 90% confidence intervals, for effect sizes based on standardised means (d , r) 95% confidence intervals are used in case of two-sided tests and 90% confidence intervals when one-sided directional tests are used (Steiger, 2004). All confidence intervals have been calculated using the R packages *MBESS* (version 4.4.0) (Kelley, 2017) and *psychometric* (version 2.2) (Fletcher, 2010).

Part II

Experimental Studies

2 Controversy Awareness on Evidence-led Discussions as Guidance for Students in Wiki-based Learning

Wikis mainly distribute user-generated content over the article and its corresponding talk page. While educational research provides article-related suggestions for learner's support, research has rarely analysed the potentials of supporting learning-related processes at the talk page level. With the presented experiment, this issue was addressed by investigating effects of visual controversy awareness information on content-related discussion threads. Such information can induce socio-cognitive conflicts which research assumes to be beneficial for learning, particularly when contradictory evidence leads wiki discussions. It was investigated how controversy awareness highlight as implicit guidance directs students' ($N = 81$) navigation and learning processes as well as their internalised knowledge representations. Results indicate that the implementation of controversy awareness representations helped students to focus on selecting meaningful discussion threads. The findings suggest that wiki talk page users can benefit from additional structuring aids and increase their learning outcome when being aware of occurring controversies.

2.1 Method

2.1.1 Design and Participants

For this experimental study, a multivariate between-subjects design was used to investigate the potential benefits of increased controversy awareness on individual learning. The single factor representing differing levels of cognitive group awareness support on controversial discussion topics, was subdivided into three levels (no additional information vs. controversy information vs. controversy + status information). The underlying Wikipedia-like structure of the talk page and discussion thread contents was identical for all three experimental groups. Participating students completed the study's tasks individually at all stages in their own wiki instance (Figure 2.1).

Artikel:"Das Aussterben der Dinosaurier" - Bearbeiten	Inhaltsverzeichnis der Diskussionen
<p>Ein Eintrag extraterrestrischen Materials durch einen Meteoriteneinschlag auf die Erdoberfläche zum erdgeschichtlichen Zeitpunkt der Kreide-Tertiär-Grenze (vor etwa 66 Millionen Jahren) können durch Forschungsergebnisse belegt werden. Die Frage, ob dieses Ereignis tatsächlich für das massenhafte Aussterben (Extinktion) der Dinosaurier verantwortlich ist, ist noch nicht restlos geklärt.</p> <p>Als möglicher Ort des Einschlags wird der Chicxulub-Krater im Golf von Mexiko nahe der Halbinsel Yucatán genannt. Kontrovers diskutierte Untersuchungen an Bohrkernen aus dem Kratergebiet durch G. Keller (2004) und M. Harting (2004) deuten allerdings darauf hin, dass der Chicxulub-Krater etwa 300.000 Jahre älter sein könnte. Dieser Einschlag könnte in diesem Fall nicht das Massenaussterben verursacht haben, das im Falle eines global verheerenden Einschlags in einem wesentlich kürzeren Zeitraum als den obigen 300.000 Jahren abgelaufen sein muss.</p> <p>Ein wesentliches Indiz für die Hypothese eines Einschlags ist der ungewöhnlich hohe Iridium-Gehalt vieler Gesteine nahe der Kreide-Tertiär-Grenze. Da der Erdmantel im Vergleich zu Steinmeteoriten arm an Iridium ist, vermutet man, dass sich in diesen Schichten der beim Einschlag aufgewirbelte Staub wiederfindet. Starke Unterstützung erhält die Hypothese eines Meteoriteneinschlags durch eine Anomalie der Chrom-Isotopenverteilung in derselben Schicht, die auch die Iridium-Anomalie enthält.</p> <p>Weitere mineralogische Spuren des Einschlages bestehen aus Ergebnissen der Druckwelle und der hohen Temperaturen, wie veränderte Quartzstrukturen, Zirkon, Diamantkristallen und Glaskugeln. Diese Strukturen kommen weltweit vor und nehmen quantitativ proportional mit der Entfernung vom Krater ab.</p>	<ul style="list-style-type: none"> Einschlagsbild(er) Verwendung des Begriffs KT-Impakt Zeitpunkt des ersten Auftretens der Dinosaurier ist zu früh Klassifikation Alternative Theorien Größter gefiederter Dinosaurier "Playing God" "Neuere" Theorien über den Meteoriteneinschlag Theorie und Praxis Alvarez-Hypothese Bindestrich oder Gedankenstrich? Ausgeglichenheit des Artikels Gut gemacht! Iridium im Erdreich ist ein starker Beleg für einen Meteoriten Das hauptsächliche Impakt-Bild Das "Datum" der KT-Extinktion ist falsch WikiProjekt Wissenschafts-Richtlinien Ceratopia oder Ceratopsia? Letztendlich sind die Dinosaurier durch Staub ausgestorben Literaturhinweise Unverständlicher Inhalt Verschieben auf Kreide-Paläogen-Grenze Defekte Weblinks Massenaussterben der Dinosaurier oder Pseudoextinktion?

Figure 2.1. Excerpt of the study's generated talk page, depicting a discussion thread comprising a resolved evidence-led controversy between two discussants in one of the experimental groups.

As dependent variables, individual learning outcomes, contribution metrics regarding article edits and discussion replies, and log data recorded by the wiki learning environment were measured. An experimentally controlled setting with individuals was used to isolate potential effects of the deployed experimental variations of the wiki instances. Social interactions that normally occur in such a setting would potentially have led to noise caused by social interactions and diversion from the intended guidance effects of controversy awareness indication. Thus, with a controlled experiment it was possible to gather evidence about potential cause-and-effect relationship if changes in the independent variable caused changes in the dependent variables. A total number of $N = 81$ students took part in this experiment, mainly recruited from the Applied Cognitive and Media Science program (67.90%) at the University of Duisburg-Essen (Germany). The students' age range was between 18 and 30 years ($M = 21.70$; $SD = 2.76$; 58 women; 23 men). All participants were assigned randomly to one of three experimental groups, resulting in an equal distribution of 27 participants per group. Due to one-time technical issues, resulting in a few missing data points, some statistical analyses had to be performed with a total number of participants of $N = 79$. At the beginning of the experiment, participants were asked to assess their interest in the experiment's topic, prior knowledge, and confidence about related hypotheses on 6-point scales ranging from "0 = low" to "5 = high" before the experiment began. The participants' overall topic-specific interest was on a medium level ($M = 2.86$, $SD = 1.20$) and their self-assessed prior knowledge about the topic was low ($M = 1.36$, $SD = 1.15$). Differences between groups in topic-specific interest were small to non-existent, $\chi^2(2, N = 81) = 0.36$, $p = .836$, $\eta_H^2 < .01$, 90% CI [.00, .03], $BF_{01} = 8.52$. Likewise, there was no meaningful difference between groups in prior knowledge, $\chi^2(2, N = 81) = 0.83$, $p = .662$, $\eta_H^2 = .01$, 90% CI [.00, .04], $BF_{01} = 7.44$. Moreover, students were also asked to assess

their confidence about the validity of the topic's related hypotheses, resulting in low to moderate confidence ratings (meteorite hypothesis: $M = 2.51$, $SD = 1.47$; volcano hypothesis: $M = 1.85$, $SD = 1.22$; other hypotheses: $M = 1.63$, $SD = 1.32$).

2.1.2 Material

The subject area that participants worked with for the entire study covered the contemporary theories of dinosaur mass extinction during the Cretaceous-Paleogene (K-Pg) boundary at around 66 million years ago. To provide a common ground for all participants in the three experimental groups, an initial base article on the topic was given to them prior to reading any of the talk page discussion threads. This article was derived from original sections of Wikipedia and adapted for the study's purpose. It had a total length of 220 words and was first presented immediately before the discussion's table of contents. It was accessible for a second time when participants had the task to edit the article with help of previously read contents on the article's corresponding discussions threads. From talk page discussions on the corresponding Wikipedia articles, a total number of twenty-four discussion threads was generated with the aim of reproducing a genuine wiki talk page environment. All threads were made up of at least two discussants and included these on a wiki talk page to represent existing discussions that directly relate to the main article. Six of the integrated discussions comprised content-related controversies with opposing points of view supported by scientific evidence on certain aspects of the presented mass extinction theory. Peripheral topics were of structural or socio-emotional nature (e.g. flames) that are also prevalent in most online discussion forums and were completely task-irrelevant distractors.

2.1.3 Wiki environment

For the experiment, a structured Wikipedia-like learning environment was created. All pages in this study were simplified versions of a Wikipedia article and its accompanying talk page, with possibly distracting items like navigation panes or logos removed. Additionally, the talk page was slightly altered from the original wiki design to enable recordings of several log data on how participants navigate through the talk page and to examine the effectiveness of the deployed controversy awareness visualisation. For that reason, the discussion threads were integrated into the page's table of contents and made them expandable and collapsible by clicking a desired thread title (Figure 2.2).

Playing God

Latest theories about the impact

By now, there is extensive evidence (Alvarez, 1986) for the fact that it is rather unlikely that the impact in front of the island of Yucatán was the crucial cause for the mass-extinction of the dinosaurs. The evidences for a luminary being the cause for this are very strong (Penfield & Camargo, 1981). But one probably searched for at the wrong place. A meteorite impact in front of India is expected to have caused the extinction of the dinosaurs. (Ebel & Grossmann, 2012). --Joe D 12:27, 24. Feb. 2014

That's just nonsense. This potential "evidences" are kind of poor. There is much more evidence for the Chicxulub crater as a probable location for the impact, see (Walther, 1978) and (Darnon, 1979). --Carbon 14 16:35, 25. Feb. 2014

Yes, this might be true what is written there. But as you can read here (Alvarez, 1986), it is not that sure what is said about Mexico. Therefore the India hypothesis is quite plausible to me. One should really mention that in the article. --Joe D 16:58, 25. Mar. 2014

I'm sure your reference isn't a bad one, but the Mexico hypothesis is well-established and is supported by a wider range of scientists. Just have a look at Darnon's overview about older and newer findings around Mexico. --Carbon 14 17:35, 26. Mar. 2014

I think this new reference is really interesting and exciting. But it seems kind of new and controversial to me. Maybe we should wait before adding it to the article? Or at least we should point out that this reference is kind of controversial. --Truthteller 17:39, 26. Mar. 2014

You're right, Truthteller. That's actually just rational. --Joe D 12:27, 27. Mar. 2014

All right, I'm okay with that, too! --Carbon 14 16:35, 29. Mar. 2014

Theory and practice

Alvarez hypothesis

Hyphens v. en dashes

Balance of the article

Well done!

Iridium in ground earth as a strong evidence for a meteorite

The main impact image

Wrong date of KT-extinction

Figure 2.2. Excerpt of the study's generated talk page, depicting a discussion thread comprising a resolved evidence-led controversy between two discussants in one of the experimental groups.

Two of the three experiment's groups were provided with additional visual information on controversial discussion topics. Both experimental groups received equal explanations about the meaning of the respective coloured indicators in a mandatory tutorial. Differences between the three groups are briefly described in the following (cf. Figure 2.3).

No highlight (Control). The representation for the non-supported group was inspired by a standard wiki talk page and did not provide any further information on the contents of a discussion thread on first sight other than its title.

Controversy highlight. For the first experimental group a supplementary external representation was added. A single-coloured indicator highlighted discussion threads on the talk page that comprised meaningful and relevant conflicting points of view that were primarily led by scientific evidence.

Controversy + Status highlight. In the second experimental group the external representation was subdivided into two classes of conflict. Two indicator colours informed participants whether a discussion thread contained unresolved conflicting points of view (red) or an already resolved conflict or consensus (green).



Figure 2.3. Illustration of a talk page excerpt for all groups: no highlight (top) vs. controversy highlight (middle) vs. controversy + status highlight (bottom).

2.1.4 Measurements

Measuring learning success. A post-experimental multiple-choice knowledge test was developed To measure individual learning success on the study's topic of dinosaur extinction events. Such tests are still widely used to quantify learning results in collaborative and individual settings (Kent, Laslo, & Rafaeli, 2016). In total the test comprised fifteen questions on the subject matter. Three of those questions were answerable with only the information provided in the original base article and therefore were solvable without having read any of the discussion threads. The remaining twelve questions were constructed in a way that each of the existing six controversial discussion topics were covered by exactly two different questions. Every question had four answering options comprised of three distractors and one attractor (Table 2.1). For the subsequent analyses, the test's overall sum was used as a general indication of individual learning success about the topic. For further detailed analyses, subscores were generated for article and discussion-related questions. The discussion-related questions were further subdivided into scores for questions matching unresolved or resolved controversial discussions.

Measuring selection behaviour. The study's modified wiki environment was able to record a participant's topic selection behaviour by measuring individual clicks on a discussion's thread title. Clicking on a title was necessary to select and expand a thread and thus unveil its contents. By design only one topic could be open for reading at a time and had to be collapsed by clicking again before proceeding to the next topic of interest. For further processing click counts that triggered only the expanding/opening events were recorded in the log.

Table 2.1

Example questions/statements and answers of the 15-item multiple-choice test.

Sample question or statement	Answer options	Relates to
The fact that birds descended from dinosaurs is an indication that...	(a) ... a mass-extinction event could not have happened. (b) ... birds cannot directly be descendants of dinosaurs. (c) ... the scientific evidence is inconclusive with regard to the origins. (d) ... only pterosaurs could have survived any event of mass-extinction.	unresolved controversy
A meteorite impact that triggered a nuclear winter would have most likely caused...	(a) ... that immediately afterwards plant eating dinosaurs (herbivores) were extinct. (b) ... that immediately afterwards meat eating dinosaurs (carnivores) were extinct. (c) ... that immediately afterwards all plants were extinct. (d) ... secondary effects that could have been responsible for the final extinction.	resolved controversy
The following mineralogical trace is NOT an indicator for an impact:	(a) iridium (b) glass (c) platinum (d) diamonds	original article

Note. Answer options in bold print represents the question's attractors.

Measuring reading times. Discussion reading times were measured by calculating the differences between opening and closing times. If a topic was opened and closed more than once, the environment's logging system also recorded cumulative reading times for each discussion thread.

Measuring sequential patterns. For each opened topic the title and index number of the previously selected topic was recorded and prepared to perform sequential pattern analyses. If no topic was preceding, it was marked as the initial topic that a participant has selected first and it was marked as the start of a sequence. For each experimental group sequence databases were created to perform analyses for discovering all frequent closed sequential patterns using the CM-ClaSP algorithm (Fournier-Viger, Gomariz, Campos, & Thomas, 2014).

Learning-related cognitive variables. Need for Cognitive Closure was measured with the 16-NCCS (Schlink & Walther, 2007). This validated questionnaire comprises 16 statements (e.g. "I prefer tasks that precisely define what needs to be done and how it has to be done.") that participants had to rate on a 6-point scale ranging from fully disagree to fully agree. Epistemic curiosity was measured with the Epistemic Curiosity Scale (Renner, 2006). This validated questionnaire measures with a total of 10 items diversive and specific epistemic curiosity. Each subscale consists of 5 statements (e.g. "When I learn something new, I like to learn even more about it.") and were rated on a 4-point scale ranging from fully disagree to fully agree.

Quality of individual contributions. Regarding the quality of knowledge artefacts that were externalised by participants into the learning environment in terms of replies and article edits, results of preliminary analyses have been published indicating promising direct and indirect effects of the provided visual guidance on

controversy awareness (Heimbuch & Bodemer, 2016a). These analyses presented first evidence for a causal relationship between the experimental variation, selection of relevant topics, reading intensity, replying preference and contribution quality of edits and discussion replies. Qualitative content analyses were conducted (Mayring, 2014; Kuckartz, 2014) on the final edits of the individual wiki articles and on each of the discussion replies. The finally derived coding schemes and category formations originated both from deductive and inductive procedures. Wiki article categories have been assigned deductively in accordance with Wikipedia's guidelines for evaluating an article's quality. Categories regarding a discussion reply's content quality have been formed in an inductive procedure. The quality of article contributions and discussion replies was measured by means of qualitative content analyses by coding and categorizing contents. These qualitative measures were also complemented with quantitative non-parametric correlation analyses and serial mediation models as path analyses where all relevant process variables were included in the statistical model.

2.1.5 Procedure

The experiment was conducted as an individual setup with up to four participants at the same time, separated by divider panels. Prior to their wiki tasks students completed a mandatory tutorial about the environment's usage and were informed about the study's goals, i.e. to learn more about the subject matter and help to (re-)construct existing knowledge artefacts by contributing to one's personal instance of a wiki article and discussion page snapshot. The tutorial introduced to them the special structure of the wiki-like environment with its distinct layers and had an exploration phase to familiarise themselves with the environment, especially with the structure of the modified talk page. In addition to that, both experimental groups

received information about the meaning of the coloured indicators corresponding to their group condition, without further explicit instructions of if or how they should use these indicators. Before the experimental phase started all groups received the same instructions about the following pages and associated tasks. The control group and both experimental groups had the same task of editing an initial Wikipedia-like base article about the mass extinction of dinosaurs and participating in up to three of the existing discussions. Since they did not receive any additional material regarding the topic, participants received the instructions that the contents of the discussions contain sufficient material to enrich the original article. So, participants could find the whole relevant information basis to fulfil the editing task on the corresponding article talk page with its twenty-four discussion threads that was generated as the additional knowledge base. At all experimental stages participants had mandatory time limits. They had maximum of five minutes for an initial read of the wiki article and a maximum of ten minutes for selecting and reading threads on the corresponding discussion page. After reading the article and discussions participants had to contribute to the wiki by first replying to three discussion threads of choice. Participants did not receive any further instructions on what kind of reply they should make for enabling them to solve or start a controversy or any other kind of possible off-task reaction that typically can be found in online discussions. Subsequently, participants were asked to work on the original article supported by the contents of the previously read discussions. In the contribution phase, they had more loose time limits for editing the article and contributing to the discussions compared to the initial reading stage. The same time limits as in the reading phase before were suggested and visualised, but instead of forcing them to the next the wiki environment automatically prompted them when time was up and kindly asked to finish the task as soon as possible.

Followed by the contribution and revision stage, the questionnaires to determine the individual levels of Need for Cognitive Closure (16-NCCS) and epistemic curiosity (ECS) were presented. After filling out these questionnaires participants had to answer a multiple-choice test about the study's contents. Finally, as an additional manipulation check, participants were asked to shortly sum up in open text fields why they have selected certain discussions to comment on and what led to the final decisions for the resulting article edits (cf. Figure 2.4).

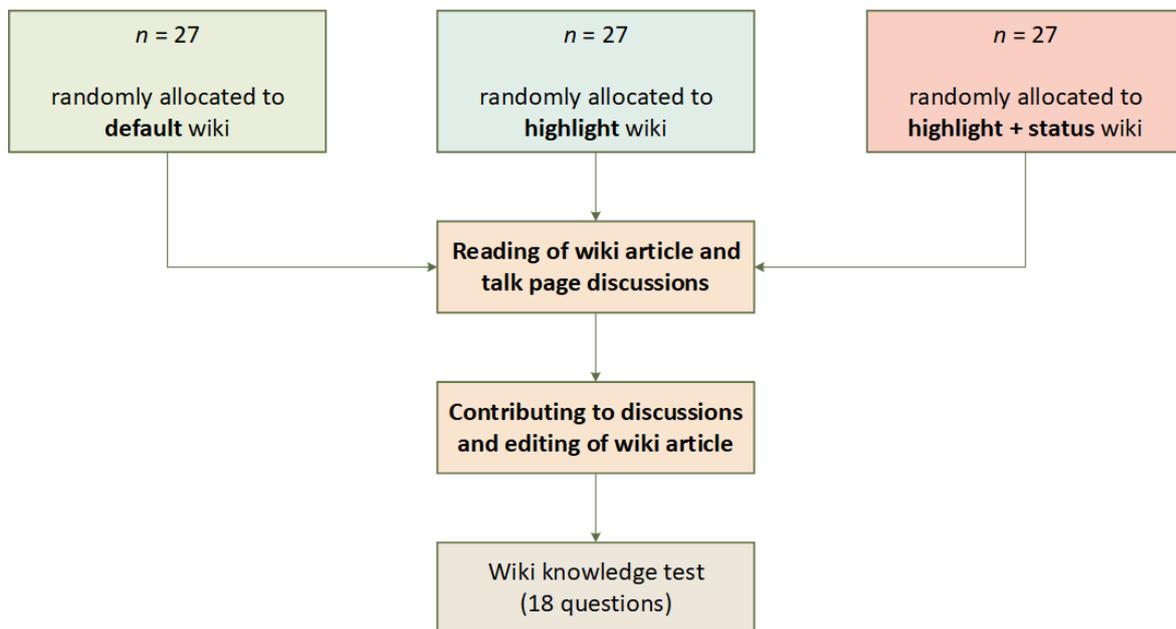


Figure 2.4. Workflow diagram visualising the overall study procedure of Experiment 1 with its central stages.

2.1.6 Specific Hypotheses to Research Questions

RQ2.1: To what extent does visual controversy awareness information influence individually relevant learning-related processes in wiki-based environments?

H2.1a: It is expected that by providing awareness information, learners will make a more focused selection of potentially important topics.

H2.1b: It is also expected that by providing awareness information, learners will read potentially most relevant topics more intensively.

H2.1c: Furthermore, It is expected that by providing awareness information about the most relevant topics, learners will make more extensive contributions to the wiki.

RQ2.2: To what extent do learners use visual controversy awareness information on evidence-led discussion threads to foster quantifiable learning outcomes?

H2.2: It is expected that providing controversy awareness information on the presence of meaningful controversies will lead to a preferred selection of relevant topics and consequently result in a higher learning outcome.

RQ2.3: To what extent do one's personal Need for Cognitive Closure and epistemic curiosity affect knowledge construction processes?

H2.3a: It is expected that individuals who are high in Need for Cognitive Closure and are provided with controversy awareness information and the status of the controversy's resolution will show a preference for resolved conflicts and show avoidance behaviour for ongoing, unresolved controversies.

H2.3b: It is also expected that individuals with high epistemic curiosity and who are provided with controversy awareness information will seek more information in peripheral topics to find new stimuli.

2.2 Results

In the following chapters, descriptive and inferential findings of the participants' activities at different stages of the experimental wiki sequence are presented by analysing the recorded data for reading and editing the talk page discussions and the corresponding article.

Discussion topic selections. The discussion thread selection preferences of participants were analysed based on recorded click events in the learning environment. Each participating student selected on average $M = 11.86$ ($SD = 5.27$) discussion threads. Analysis of variance using planned comparisons with an orthogonal Helmert contrast indicates that in the experiment's unsupported control group, students selected on average three discussion threads more on the article's talk page compared to both supported conditions, $F(2, 78) = 3.80$, $p = .027$, $\eta^2 = .09$, 90% CI [.01, .18], $BF_{10} = 2.02$. The first comparison between the control and both experimental groups showed a meaningful difference (Contrast Estimate = 2.76, $SE = 1.20$, $p = .024$), whereas the contrast estimate for the comparison between the two supported groups differed less (Contrast Estimate = 2.11, $SE = 1.39$, $p = .132$). From all groups, participants in the group with additional controversy status information where the most focused by selecting the lowest number of discussion threads for reading.

To further analyse these inferential findings on selection behaviour in the different experimental groups, a sequential pattern mining analysis was performed with SPMF using the CM-ClaSP algorithm (Fournier-Viger et al., 2014). Participants who did not receive additional awareness information on the discussion type showed a tendency to follow a top-down reading strategy, beginning by reading the very first discussion thread on the experimental wiki talk page. In contrast, members of

both supported awareness visualisation groups primarily focused on selecting the most relevant topics first. Both quantitative and qualitative analyses support hypothesis H2.1a on a more selective behaviour in either visually supported group (cf. Table 2.2).

Table 2.2

Most frequent closed sequential patterns discovered by the CM-ClaSP algorithm.

Group	Thread sequence	Pattern frequency
No highlight (control)	1, 2, 3, 5 , 7	14
Controversy highlight	5, 8, 14, 16, 19	14
+ Status highlight	5, 8, 14	16
	8, 14, 19	15
	5, 8, 19	13

Note. Content-related thread numbers are highlighted in bold print.

In addition to the closed sequential pattern analysis, state transition probabilities with markov models were calculated for the three studied groups (cf. Figure 2.5). Participants in the control group without additional highlights always started to read peripheral threads first, with a likelihood of 1.0. Not a single participant has selected a controversial topic as an initial thread to read, not even by chance through randomly clicking on a topic title. For the first experimental group, when a single-coloured (blue indicator) controversy highlight was added to the discussion thread titles the likelihood of starting with a peripheral thread was lowered to .67. For the second experimental group with two-coloured (green and red indicators) controversy status highlights the likelihood of starting with a peripheral thread was lowered even more to .42. Participants in both experimental groups were attracted also towards controversial discussion topics as initial choice.

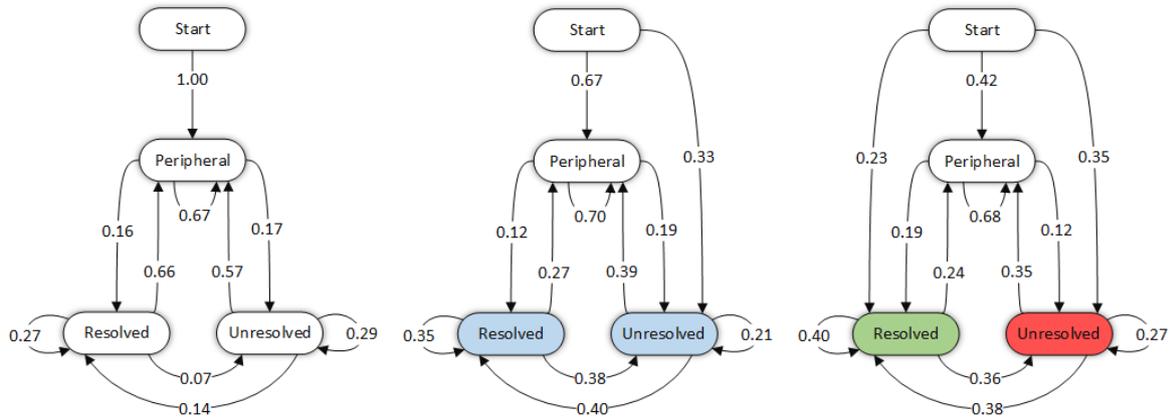


Figure 2.5. Markov models with state transition probabilities for *no highlight (control)* (left), *controversy highlight (middle)* and *+ status highlight (right)*.

Wiki reading times. Overall, each participant spent on average $M = 229.01$ seconds ($SD = 42.52$) on reading the original wiki article and $M = 488.37$ seconds ($SD = 73.13$) on reading the corresponding talk page. Regarding the time spent on the entire discussion page, there was a rather small difference between the three groups, $F(2, 70.72) = 1.13$, $p = .330$, $\eta^2 = .03$, 90% CI [.00, .09], $BF_{01} = 3.87$. More detailed analyses of categorised reading times relative to the entire discussion page reception suggest several differences (cf Table 2.3). As one of the study's main interests, relative reading times were analysed focusing on resolved and unresolved controversies using Helmert contrasts. There was a substantial difference between the groups with a large effect, $F(2, 71.43) = 21.08$, $p < .001$, $\eta^2 = .35$, 90% CI [.20, .46], $\log BF_{10} = 12.48$. The first comparison between the unsupported group and both supported groups showed a substantial difference (Contrast Estimate = -0.29, $SE = 0.05$, $p < .001$), whereas the contrast estimate for the comparison between both supported groups indicates only a negligible difference (Contrast Estimate = -0.05, $SE = 0.05$, $p = .374$).

Table 2.3

In-depth analyses on the relative reading times of the different discussion thread categories.

Group	Category	<i>M (SD)</i>	<i>F(2, 78)</i>	<i>p</i>	η^2	90% CI
No highlight	Resolved controversies	.19 (.12)	14.88	< .001	.28	[.13, .39]
	Unresolved controversies	.19 (.07)				
	Peripheral	.53 (.17)				
Controversy highlight	Resolved controversies	.33 (.15)	8.69	< .001	.18	[.06, .29]
	Unresolved controversies	.32 (.12)				
	Peripheral	.25 (.22)				
+ Status highlight	Resolved controversies	.41 (.18)	21.06	< .001	.35	[.20, .46]
	Unresolved controversies	.29 (.15)				
	Peripheral	.21 (.19)				

Analyses of pooled absolute reading times on the study's relevant controversial topics showed differences in individual thread reading and selection behaviour dependent on the discussion category and whether the participant received additional awareness information or not. Participants spent less time on reading topics comprising any kind of temporary consensus when no additional awareness visualisation was provided ($M = 103.74$, $SD = 60.21$) compared to the controversy highlight groups ($M = 196.26$, $SD = 84.96$), $t(69.66) = 5.65$, $p < .001$, $d = 1.26$, 90% CI [0.90, 1.76], $\log BF_{10} = 9.29$. Accordingly, there was a difference in reading times of unresolved controversial discussions between participants of the control group ($M = 102.48$, $SD = 38.32$) and both of the controversy highlight groups ($M = 161.06$, $SD = 75.10$), $t(78.94) = 4.65$, $p < .001$, $d = 0.98$, 90% [0.68, 1.51], $\log BF_{10} = 5.20$. Conversely, the

data suggests that members in the control group ($M = 292.37$, $SD = 107.83$) read peripheral discussions for longer time periods than participants in the controversy highlight groups ($M = 125.94$, $SD = 113.90$), $t(54.74) = 6.43$, $p < .001$, $d = 1.50$, 90% CI [1.05, 1.96], $\log BF_{10} = 14.05$. These differences in the data (cf. Figure 2.6) are surprisingly large assuming a true point null hypothesis model. Thus, these findings are regarded as support for hypothesis H2.1b on more intensive reading of relevant topics through increased awareness.

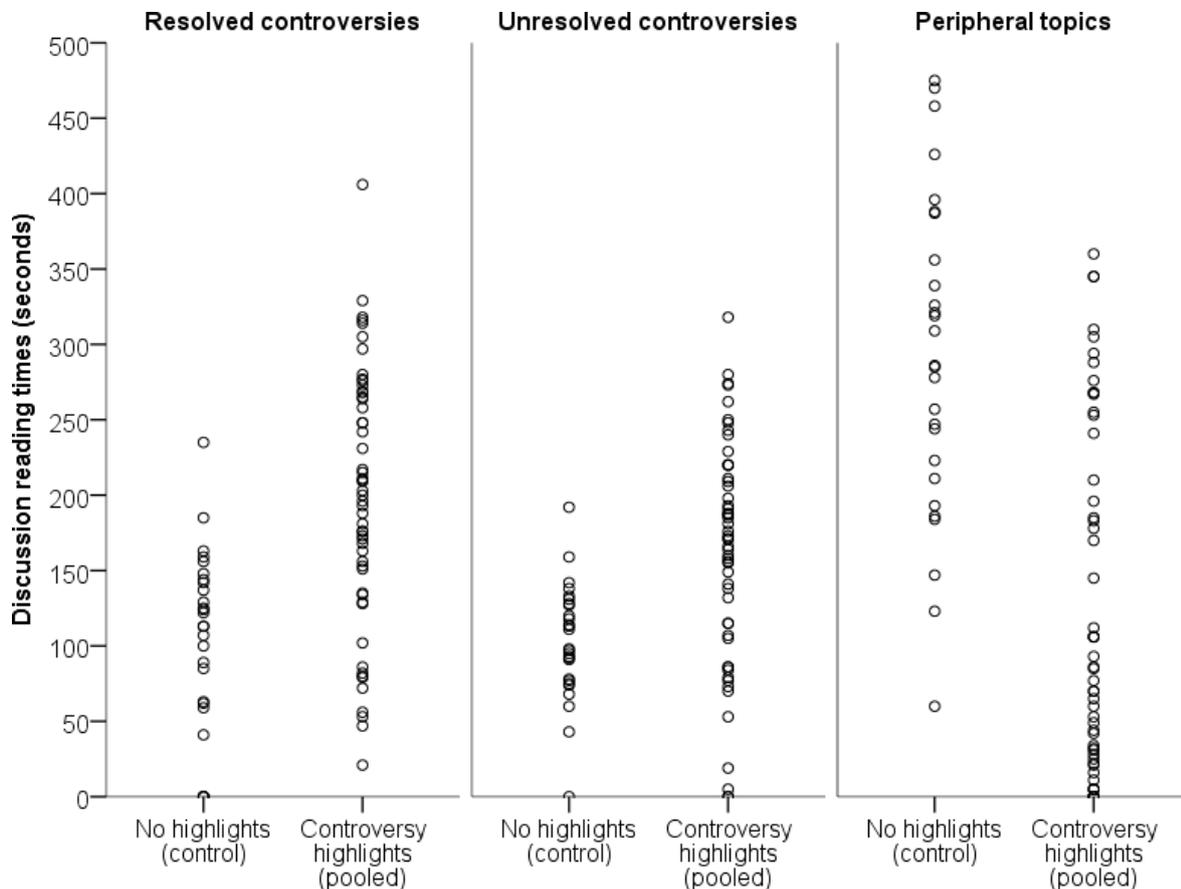


Figure 2.6. Distributions of absolute reading times in seconds for the discussion categories.

Wiki contributions. Participants' tasks in this study were to comment on discussion threads on the talk page and to revise the original article after reading selected discussion threads. The final article revision lengths ranged from 218 to 434 words (original article: 220 words), with an overall average length of $M = 279.89$ ($SD = 37.13$). Control group members wrote the shortest article with an average of $M = 274.88$ ($SD = 27.39$) words. On a descriptive level, the more detailed awareness highlight participants received, the more additions to the final article were made, with an average of $M = 279.33$ ($SD = 30.30$) words in the group with controversy highlight and respectively an average article length of $M = 285.46$ ($SD = 50.37$) in the group receiving additional controversy status information. A one-way ANOVA did not indicate substantial differences between the three groups, $F(2, 56.36) = 0.52$, $p = .596$, $\eta^2 < .01$, 90% CI [.00, .06], $BF_{01} = 6.06$. Hypothesis H2.1c on article length differences between the groups cannot be supported by these results. Analysing the amount of time participants took to write comments on up to three self-chosen discussions using planned comparisons with an orthogonal Helmert contrast, there was a moderate effect, $F(2, 78) = 2.38$, $p = .099$, $\eta^2 = .06$, 90% CI [.00, .14], $BF_{01} = 1.47$. The first comparison between the unsupported group and both supported groups showed a difference (Contrast Estimate = -61.41, $SE = 29.54$, $p = .041$), whereas the contrast estimate for the comparison between the two supported groups did not differ substantially (Contrast Estimate = 22.67, $SE = 34.12$, $p = .508$). With regard to replying behaviour, multivariate Helmert contrasts indicate a large main effect for the groups, $\lambda = 0.67$, $F(6, 150) = 5.48$, $p < .001$, $\eta_p^2 = .17$, 90% CI [.06, .24]. Detailed contrast comparisons are shown in Table 2.4. Participants without visual controversy awareness information replied more frequently to peripheral discussions instead of relevant controversial discussions, $F(2, 77) = 18.06$, $p < .001$, $\eta_p^2 = .32$, 90% CI [.17, .43], $\log BF_{10} = 10.59$.

Table 2.4

Multivariate Helmert contrast comparisons on replying to discussion categories.

Contrasted groups	Replying to	Contrast Estimate (SE)	p
No highlight vs. Controversy highlight	Resolved	-0.74 (0.20)	<.001
	Unresolved	-0.34 (0.20)	.084
	Peripheral	1.19 (0.21)	<.001
Controversy vs. + Status highlight	Resolved	0.08 (0.23)	.722
	Unresolved	0.14 (0.22)	.543
	Peripheral	-0.24 (0.24)	.315

Discussion reply quality. By using a tertile split on the categorised discussion reply data, over all experimental groups and participants, 30.86% of all participant contributions were rated as lower quality, 58.02% as medium quality and 11.11% as higher quality contributions. In Table 2.5 the aggregated quality ratings are presented for reply contributions that are partitioned according to the experimental groups. These numbers are total aggregates of inductively formed categories associated with discussion reply quality. Reply quality was composed of *Neutrality* (acknowledging both opposing arguments), *New aspects* (introducing new knowledge artefacts) and *Summarising* (integrating summaries of preceding discussion).

Table 2.5

Frequencies (and percentages) of cumulative overall quality ratings of discussion replies.

Overall quality	No highlight	Controversy highlight	+ Status highlight
Lower	12 (44.44%)	5 (18.52%)	8 (29.63%)
Medium	13 (48.15%)	19 (70.37%)	15 (55.56%)
Higher	2 (7.41%)	3 (11.11%)	4 (14.81%)

Furthermore, correlations between three inductively formed reply quality categories and the replies' overall quality have been calculated. For each correlation between a category and overall quality the effects of the corresponding category have been subtracted from the overall quality rating. For the category of *New aspects* there was no measurable effect on the overall discussion reply quality, $\rho(79) = -.03$, $p = .828$, BCa 95% CI [-.25, .19], $BF_{01} = 6.96$. The positive effect of categorised *Neutrality* on quality was relatively small, $\rho(79) = .20$, $p = .073$, BCa 95% CI [-.08, .45], $BF_{01} = 1.49$. Between the category *Summarising* and reply quality, there was a moderate to strong positive effect, $\rho(79) = .39$, $p < .001$, BCa 95% CI [.14, .59], $BF_{10} = 80.32$. In addition to the analyses of categories, there was also a moderate to large positive correlation between the reply length and the overall quality, $\rho(79) = .47$, $p < .001$, BCa 95% CI [.25, .64], $\log BF_{10} = 7.65$. In-depth examinations of the effects of the category *Summarising* on the contribution quality of discussion replies with regard to the topic selection behaviour, suggest moderate positive effects for *Replying to Controversial Discussions* ($\rho(79) = .31$, $p = .010$, BCa 95% CI [.06, .48], $BF_{10} = 6.83$) and a small to moderate negative effect for *Replying to Peripheral Discussions* ($\rho(79) = -.26$, $p = .021$, BCa 95% CI [-.45, -.02], $BF_{10} = 2.06$). Additional statistical frequency analyses for this category can be found in Table 2.6.

Table 2.6

Frequencies (and percentages) of integrating summaries of the preceding discussion into replies.

Summarising	No highlight	Controversy highlight	+ Status highlight
No	8 (29.63%)	4 (14.82%)	7 (25.93%)
Abstract	3 (11.11%)	1 (3.70%)	0
Elaborate	16 (59.26%)	22 (81.48%)	20 (74.07%)

A path analysis model in form of a multiple serial mediation was used to analyse the relevant process variables of the experiment. The group effects on discussion reply quality were analysed with the relevant process variables *controversy selection*, *controversy reading time* and *replying to controversies* as mediators. There was an indirect effect of the controversy awareness group on reply quality, if users followed the whole process of selecting, reading and replying to controversial discussions, $Effect = .07, SE = .04, 95\% CI [.01, .15]$. Users who received additional controversy awareness information via highlights were more likely to select controversial discussions, consequently they spend more time on reading controversies, followed by an increased likelihood to reply more frequently to controversies and finally resulting in higher quality discussion replies (cf. Figure 2.7).

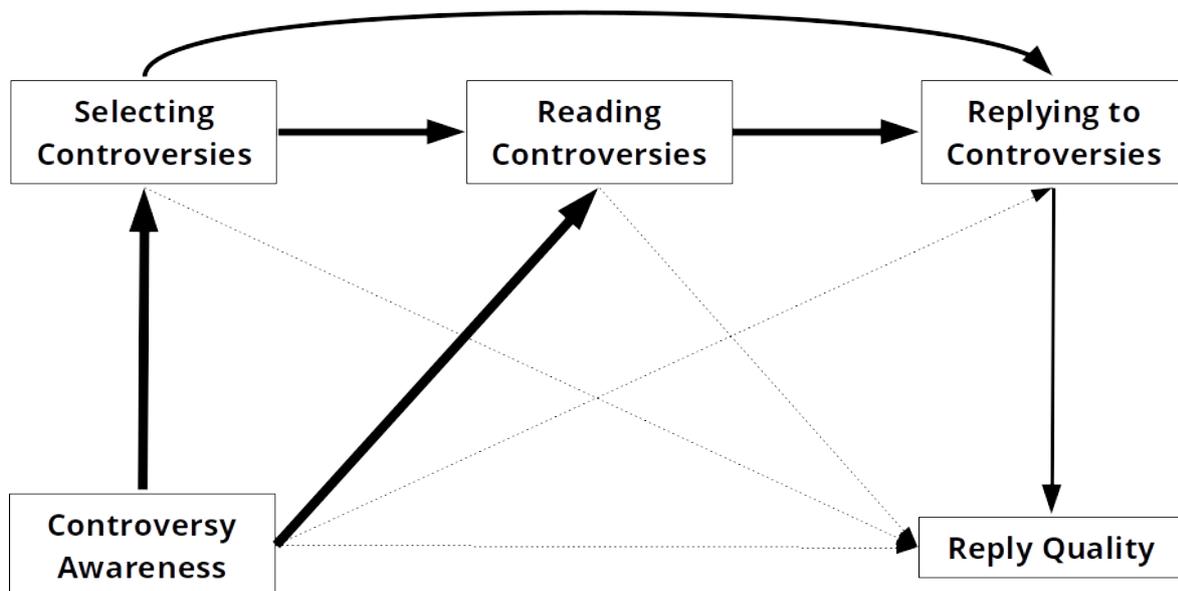


Figure 2.7. Path analysis on reply quality with process variables (selecting, reading, replying) as serial mediators. Solid lines represent regressions below an $\alpha = .05$ threshold, with bolder lines representing lower p -values.

Article edit quality. Over all groups and participants, 16.46% of all final articles were rated as lower quality, 51.90% as medium quality and 31.65% as higher quality edits. In Table 2.7 the aggregated article quality ratings are presented that are partitioned according to the experimental groups. These numbers are total aggregates of deductively assigned categories based on Wikipedia’s article quality evaluation guidelines. Article quality was composed of *Referencing* (using references for article changes), *Structure* (fitting into the article’s structure), *Neutrality* (integrating neutral points of view) and *Relevance* (editing is meaningful for the article).

Table 2.7

Frequencies (and percentages) of cumulative overall quality ratings of edited articles.

Overall quality	No highlight	Controversy highlight	+ Status highlight
Lower	5 (19.23%)	2 (7.41%)	6 (23.08%)
Medium	15 (57.09%)	16 (59.26%)	10 (38.46%)
Higher	6 (23.08%)	9 (33.33%)	10 (38.46%)

Correlations were calculated between deductively assigned categories and overall quality of the resulting articles to gain deeper understanding what types of edits to an article add the most value. The effects of each category regarding article quality have been subtracted from the appropriate overall quality ratings. *Referencing* as qualitative category is positively associated with the overall ratings of article quality with a small effect, $\rho(77) = .22$, $p = .048$, BCa 95% CI [.70, .94], $BF_{01} = 1.10$. For the relation between *Structure* and article quality, a moderate to large positive effect has been found, $\rho(77) = .43$, $p < .001$, BCa 95% CI [.22, .58], $\log BF_{10} = 5.72$. The effect for *Neutrality* was large and positively correlated with overall quality, $\rho(77) = .53$, $p < .001$, BCa 95% CI [.34, .68], $\log BF_{10} = 10.42$. Between the category *Relevance* and

edited article quality, there was a large effect that was positively related to the overall ratings, $\rho(77) = .56, p < .001$, BCa 95% CI [.41, .69], $\log BF_{10} = 12.17$. In addition to the analyses of categories, there was also moderate positive correlation between the editing amount (added or changed characters) and the edited article quality, $\rho(77) = .35, p = .002$, BCa 95% CI [.08, .57], $BF_{10} = 74.66$. As before where analyses on the discussion reply quality were presented, a corresponding path analysis model was used in form of a multiple serial mediation for group effects on article edit quality with the same process variables as mediators. In contrast to the discussion reply quality model, the indirect effect of the controversy awareness information on edit quality did not fully transfer to the final article, $Effect = .02, SE = .03, 95\% CI [-.03, .09]$ (cf. Figure 2.8).

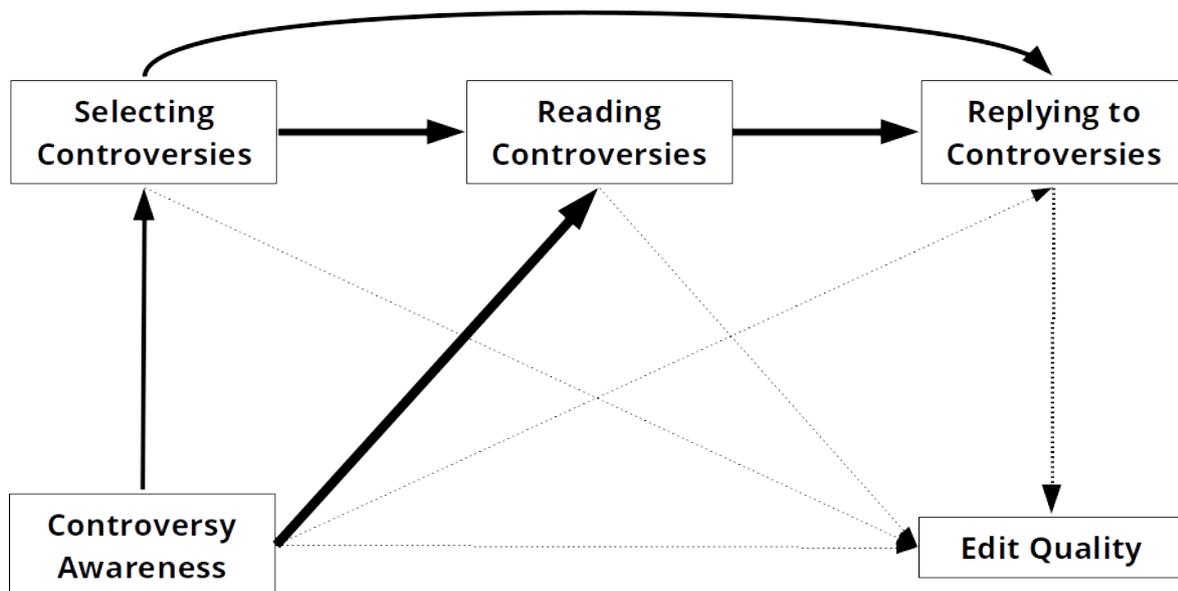


Figure 2.8. Path analysis on article edit quality with process variables (selecting, reading, replying) as serial mediators. Solid lines represent regressions below an $\alpha = .05$ threshold, with bolder lines representing lower p -values.

Learning outcome. As one of the central study interests, it was analysed if providing any kind of controversy awareness support led to a generally positive learning outcome that should have manifested in a higher knowledge test score. On a descriptive level, all three groups showed only fractional differences with $M_1 = 9.85$ ($SD_1 = 1.90$), $M_2 = 9.78$ ($SD_2 = 2.17$) and $M_3 = 9.93$ ($SD_3 = 2.32$). Multivariate analysis of variance indicate a small to moderate effect between the three investigated groups on a global level, $\lambda = 0.93$, $F(6, 152) = 0.97$, $p = .447$, $\eta^2 = .04$, 90% CI [.00, .07], $BF_{10} = 2.80$. More details on the learning outcomes about the different categories of the questions on the multiple-choice test are presented in Table 2.8. Separate independent ANOVAs did not indicate substantial differences for any question type reference and no more than small effects.

Table 2.8

In-depth analyses on the number of correct answers in the knowledge test.

Group	Questions refer to	M (SD)	$F(2, 78)$	p	η^2	90% CI
No highlight	Original article	2.52 (0.70)	1.40	.252	.03	[.00, .11]
	Resolved	2.19 (0.83)				
	Unresolved	2.26 (0.76)				
Controversy highlight	Original article	3.85 (0.95)	0.41	.662	.01	[.00, .05]
	Resolved	3.70 (1.07)				
	Unresolved	3.96 (1.13)				
+ Status highlight	Original article	3.48 (1.28)	0.66	.520	.02	[.00, .07]
	Resolved	3.89 (1.25)				
	Unresolved	3.70 (1.38)				

In further analyses with respect to the experiment's categories of controversial discussion types (resolved vs. unresolved vs. peripheral), the test scores among the three conditions were investigated considering the categorised discussion reading times as mediators in a parallel multiple single-step mediation analysis using Model

4 within the PROCESS macro for SPSS (Hayes, 2013). The total effect model encompassed just a small overall effect, $F(1, 79) = 0.46$, $p = .501$, $R^2 = .01$, 90% CI [.00, .08], $BF_{10} = 1.32$. Nonetheless, the analysis indicates an increase in the learning outcome for those participants receiving controversy awareness information ($a_2 = 24.20$, $p < .001$) mediated by spending more time on reading unresolved controversies more intensively ($b_2 = 0.01$, $p < .001$) (Figure 2.9).

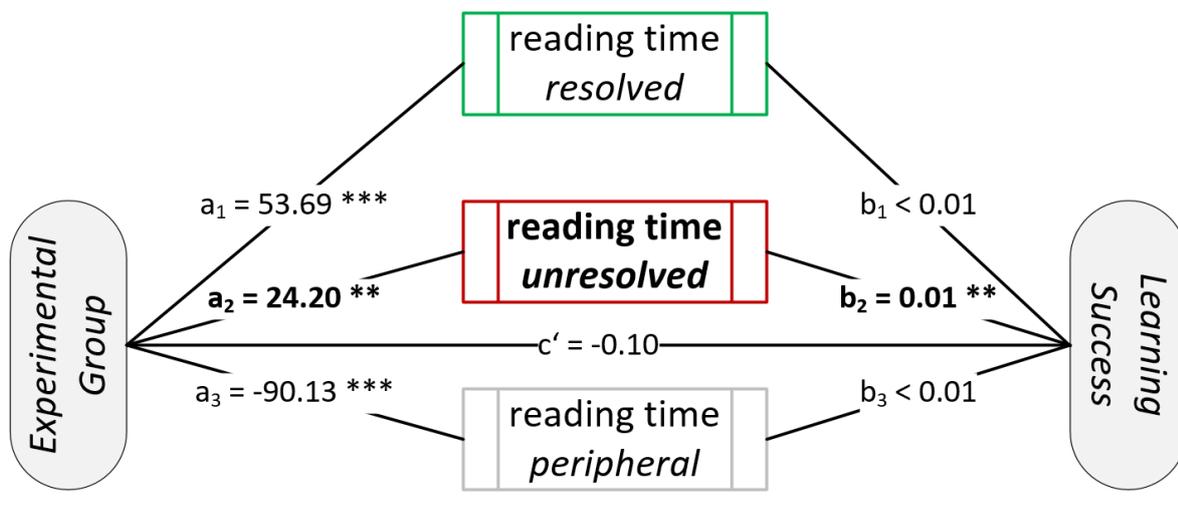


Figure 2.9. Multiple single-step mediation model on the multiple-choice test results as an outcome variable. Unstandardised a and c' weights are measured in seconds. Unstandardised b weights are measured in number of correct answers. ** $p < .01$, *** $p < .001$.

A bias-corrected bootstrap confidence interval (95% CI) for the indirect effect ($a_2b_2 = 0.21$) based on 5,000 bootstrap samples was entirely above zero (0.04 to 0.46). There was no evidence that the visual representation of controversy awareness information influenced the learning outcome independent of its effect on reading intensity of unresolved controversies ($c' = -0.10$, $p = .733$). These results in conjunction with the reported analysis of variance partly support hypothesis H2.2 of increasing individual learning outcome by providing controversy awareness indicators.

Individual cognitive influences. Furthermore, potential influences of the cognitive variables of interest were investigated, i.e. epistemic curiosity and Need for Cognitive Closure, on topic selection, reading and discussion reply frequency on unresolved or resolved controversies. For analysing the effect of Need for Cognitive Closure in interaction with the provision of visual controversy awareness support on reading times of relevant discussion threads, a one-way MANCOVA was conducted on replying behaviour to the study's discussion types as dependent variables and the Need for Cognitive Closure as covariate. The multivariate effect of the Need for Cognitive Closure on replying behaviour was moderate, $\lambda = 0.93$, $F(3, 74) = 1.89$, $p = .138$, $\eta_p^2 = .07$, 90% CI [.00, .16], $BF_{10} = 8.92$ (cf. Figure 2.10).

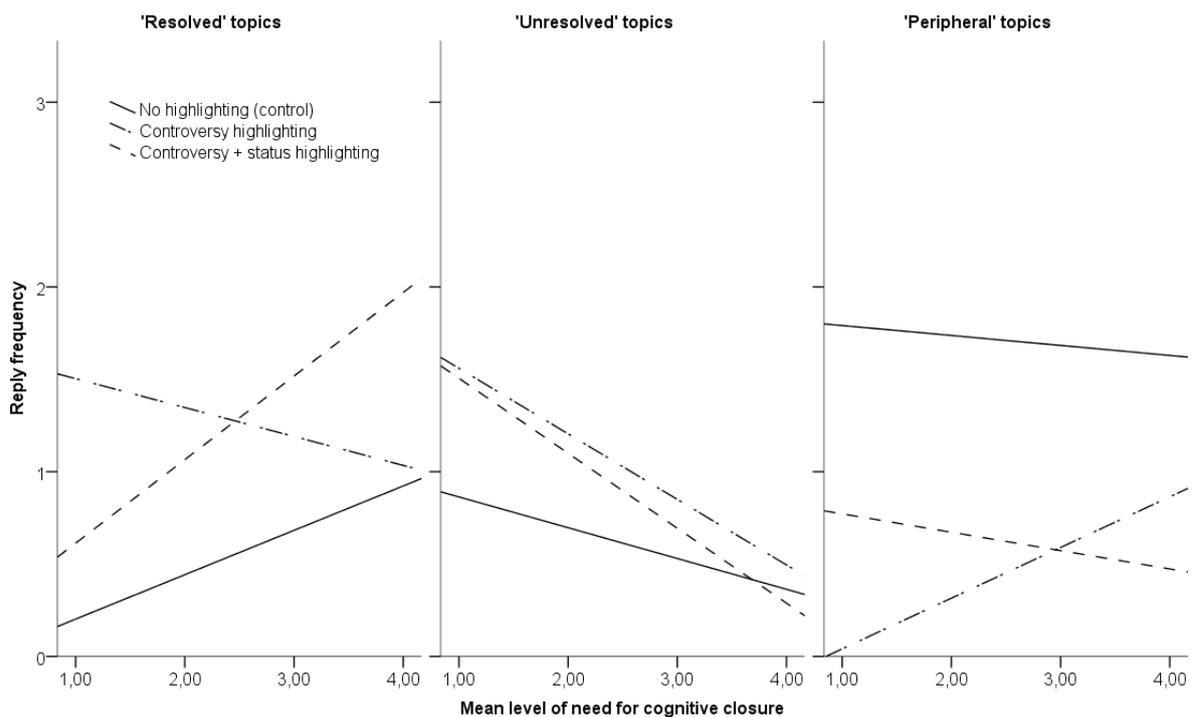


Figure 2.10. Multivariate interaction graphs between groups and the individual Need for Cognitive Closure on replies within the study's different discussion types (resolved vs. unresolved vs. peripheral).

The univariate effect of the Need for Cognitive Closure on replying to unresolved controversies was moderate, $F(1, 76) = 4.42, p = .004, \eta_p^2 = .06, 90\% \text{ CI } [.00, .15], BF_{10} = 6.97$, indicating that the replies to these topics depend on the individual Need for Cognitive Closure and the type of provided controversy awareness information. Univariate effects of the Need for Cognitive Closure on replying to resolved and peripheral topics were small and trivial with $F(1, 76) = 1.73, p = .193, \eta_p^2 = .02, 90\% \text{ CI } [.00, .10], BF_{10} = 2.15$, respectively $F(1, 76) = 0.04, p = .848, \eta_p^2 < .01, 90\% \text{ CI } [.00, .03], BF_{10} = 1.02$. For epistemic curiosity, there were several effects for individual levels of epistemic curiosity and topic selection behaviour as well as for correlations with reading times of different discussion categories. Overall, there was a small to moderate effect that the more epistemically curious a participant was, the more effort was invested in seeking additional information. First, further discussion topics were selected, $r(79) = .26, p = .010, 90\% \text{ CI } [.08, .42], BF_{10} = 3.83$. Second, more time was spent on reading additional peripheral threads, $r(79) = .21, p = .030, 95\% \text{ CI } [.03, .38], BF_{10} = 1.53$. More detailed correlations, subdivided into the study's groups of controversy awareness visualisations and discussion categories are presented in Table 2.9 for topic selection behaviour and Table 2.10 for relative discussion reading times. The results of these analyses on the influencing variables provide some evidence in favour of both hypotheses H2.3a (Need for Cognitive Closure) and H2.3b (epistemic curiosity).

Table 2.9

Correlations between epistemic curiosity and topic selections.

Group	Discussion type	<i>r</i> (79)	<i>p</i>	90% CI
No highlight	Overall	-0.41	.017	[-.58, -.21]
	Controversies	0.16	.212	[-.02, .33]
	Peripheral	-0.44	.012	[-.58, -.28]
Controversy highlight	Overall	0.52	.003	[.37, .64]
	Controversies	-0.03	.434	[-.21, .15]
	Peripheral	0.49	.005	[.34, .62]
+ Status highlight	Overall	0.59	.001	[.46, .70]
	Controversies	0.09	.327	[-.10, .27]
	Peripheral	0.53	.002	[.38, .65]

Table 2.10

Correlations between epistemic curiosity and topic reading times.

Group	Discussion type	<i>r</i> (79)	<i>p</i>	90% CI
No highlight	Resolved	-0.01	.491	[-.19, .17]
	Unresolved	-0.11	.302	[-.29, .08]
	Peripheral	0.05	.395	[-.14, .23]
Controversy highlight	Resolved	-0.45	.009	[-.59, -.29]
	Unresolved	-0.23	.121	[-.40, -.05]
	Peripheral	0.46	.008	[.30, .59]
+ Status highlight	Resolved	-0.45	.009	[-.59, -.29]
	Unresolved	0.01	.481	[-.17, .19]
	Peripheral	0.31	.057	[.13, .47]

2.3 Discussion

Knowledge construction and learning within wiki environments presents users as learners with major challenges. Talk page discussions can become exceedingly long and due to their special structure as a normal wiki page, instead of classical threaded discussions in online forums, potentially interested readers or new discussants can be overwhelmed and feel lost in the amount of potentially relevant or irrelevant information. Consequently, significant knowledge artefacts could remain completely undiscovered and unknown to an interested learner, such as deeply elaborated discussions on controversial aspects of a topic that entirely disappear in the depths of a talk page. It has been argued that providing cognitive awareness information as implicit representational guidance in the form of supportive visual indications of controversial discussion contents could lead an interested reader to a more focused selection of important knowledge artefacts. Furthermore, a positive learning outcome for learners was expected by benefiting from the reception of socio-cognitive conflicts that are positively associated within collaborative knowledge construction settings. The underlying processes that affect the effectiveness of the investigated guidance mechanisms were expected to be influenced by different inter-individual cognitive variables, that is an individual's capabilities to deal with ambiguous information and personal preferences in information-seeking behaviour regarding epistemic knowledge gains. To answer this study's first research question on the potential for guidance of controversy awareness information on the discussion thread level, the influences on the working processes in wiki-based knowledge construction were investigated. Therefore, one specific interest was the investigation whether individual participants who received any kind of additional visual structuring aids in a private wiki instance focused more effectively on relevant discussions about meaningful contents and specifically led by evidence. Analyses of variance of the selection and

reading behaviour of the experiment's topics suggested that learners were following a desired selection pattern towards the most meaningful discussions. Due to a lack of structuring aids, participants of the control group were less focused and selected more talk threads that did not contain any meaningful discussions. Thus, participants of the control group had less remaining time to seek and identify discussions that were relevant for revising the article. Therefore, they were not able to read the content-related meaningful controversies as extensively as participants of both supported groups. Indications of a more unstructured information search behaviour in the control group were additionally supported by analysing patterns of discussion topic selection. Sequential pattern mining showed that in both supported groups a focused selection of the most relevant topics occurred in contrast to a sequential reading strategy used by the control group, who had no further indications of what contents were hidden inside a topic's discussion thread. That means when wiki learners received additional awareness information about controversies, it was used by participants and worked as intended, as a structuring aid. Regarding article revision, it was expected that members in both supported groups would identify more relevant evidence to enrich the original article with new knowledge artefacts and thus produce more text, in accordance to the results found by Wichmann and Rummel (2013). While quantitative text production does not generally indicate a higher text quality, in this study's setting it indicates that learners identified more important evidence that could extend the article meaningfully. However, the data suggests only a minimal tendency on a descriptive level that more text is produced depending on the kind of support. Additionally, it was examined how the quality of contributions was affected by adding visual controversy information to relevant discussion threads. First, the immediate effects of controversy awareness information on the quality of discussion thread contributions were analysed by deductively

categorising participants' utterances. It was found that the provision of additional representations can make the occurrence of controversies more salient to users and direct them to relevant contents. This implicit guidance led to generally more elaborate contributions and overall higher reply quality. Contributions of the highest quality were marginally more frequent when the level of controversy awareness information was raised by providing additional indication on the resolution status. Further correlation analyses on the separate qualitative categories suggest the most substantial association exists between summarising relevant aspects of preceding discussions and the overall quality of their contributions. It could also be seen that when controversy awareness without additional status information was provided the most contributions of medium quality could be found, as well as the lowest number of participants who did not summarise relevant preceding discussions. In a subsequent step, indirect effects of controversy awareness representations on the quality of final article edits were analysed. Comparable to the discussion reply findings, it has been identified that the greatest number of the highest overall quality article was present in the experimental group which received the most detailed level of controversy awareness and status information. All inductively assigned categories that have been in accordance with Wikipedia's guidelines on evaluating article quality, were positively associated with the rated overall quality in this experiment whereas the strongest effect could be identified for the category Relevance. Since the most relevant additional contents to enrich the article were to be found in the relevant content-related controversial discussions, this result is in accordance with the expectations. But there was also a rather positive effect for the group that did not receive the controversy status information, that is they produced the smallest number of lower quality articles. Those findings that are in favour of the effects of the more general information on the mere occurrence of controversies can

be due to similar effects that have been found in previous wiki-related research on collaborative knowledge construction. Medium levels of incongruity of information between the individual cognitive and the socio-technical system have been proven to be the most beneficial for learning (Moskaliuk et al., 2009) as well as medium levels of redundancy fostered external accommodation processes (Moskaliuk et al., 2012). These findings are in line with research that medium levels of supportive manipulations to scaffold learning processes are beneficially for many collaborators. The results regarding quality analyses show that additional representations of controversy awareness information can be effective for the resulting socially shared artefacts, directly on the level where visualisations are implemented and even when transferred to a subsequent level. It could be shown that contributions to the wiki were of higher quality, more elaborate and relevant to the subject matter when students were provided with guidance. With the study's second research question, it was investigated whether visual awareness information on the occurrence of resolved and unresolved controversies that can be found inside wiki talk pages lead to measurable differences in the learning outcome about the subject matter. Two varying degrees of controversy awareness representations were compared with a control group that had no further information on the discussions' contents. Regarding learning outcomes, direct comparisons on a general level did not indicate differences between the three groups. Considering the different types of implemented discussions within the experimental talk page, i.e. resolved vs. unresolved controversies and peripheral topics, a parallel multiple mediation analysis identified a meaningful effect. Students benefited more in terms of learning success from an increased degree of guidance towards conflicts arisen from content-related controversies, assuming they spent more time on reading unresolved controversial discussions more extensively. However, higher reading times of resolved controversial

topics did not lead to measured knowledge gains, even though both types of controversies had an equal number of test questions. One possible explanation for not reaching conventional levels of statistical significance on the resolved controversies path in the mediation model could be that a relative imbalance in the distribution of the Need for Cognitive Closure in the current student sample, who predominantly favoured controversial opinions. Another reason for not finding an effect of the resolved controversy reception could be that presenting information associated with negativity and conflicts triggers increased cognitive activity (van Marle, Hermans, Qin, & Fernández, 2009) and elicit more arousal leading to better memorisation (Bradley & Lang, 2007) that can be explained by enhanced visual attention processes (Calvo & Lang, 2004; Nummenmaa, Hyönä, & Calvo, 2006). Nevertheless, first qualitative investigations of the individual wiki contributions showed some positive and promising trends in favour of both experimental groups (Heimbuch & Bodemer, 2016a). Students who received visual guidance on controversial discussions produced more extensive and more comprehensive discussion replies of higher quality compared to the control group. Beyond such direct effects of controversy awareness implementations on the talk page, it could also be shown that there were some positive indirect carry-over effects to the article edits. Students in both experimental groups performed more extensive and meaningful edits by including references, adding, or restructuring knowledge artefacts and keeping a neutral point of view by equally addressing opposing evidence on the subject matter. To answer this study's third research question whether cognitive variables that have been identified as relevant for learning in contexts where socio-cognitive conflicts can occur, potential influences of these variables on topic selection and replying behaviour were investigated. Therefore, specific analyses on the influential effects of Need for

Cognitive Closure were conducted, which is closely related to one's personal preference for or against ambiguity, and the extent of epistemic curiosity, which is closely related to exploratory information seeking behaviour. Students with a high Need for Cognitive Closure favoured replying to controversies that were resolved during the discussions when provided with controversy awareness visualisations. This is in total accordance with previous research that these individuals prefer to avoid ambiguous situations and information (Schlink, 2009) as well as being more impatient to come to more complex conclusions that require extensive information processing capabilities and the analysis of multiple interpretations of facts (Kruglanski & Mayseless, 1987; Kruglanski & Webster, 1996). As expected, the largest effect for the Need for Cognitive Closure impact on replying behaviour could be seen for the experimental group that was provided with visual information on the occurrence of a controversy and its status. Students with low Need for Cognitive Closure scores replied equally to resolved and unresolved controversies. When wiki learners did not receive controversy awareness support, participants replied primarily to peripheral discussions, regardless of the personal Need for Cognitive Closure. Regarding the effects of epistemic curiosity, there were small to moderate effects for selecting more additional topics as well as small to moderate effects for spending more time on reading peripheral discussions. In general, when visual controversy awareness support was provided, the higher a student scored on the epistemic curiosity scale, the more time they spent in reading additional discussion threads that were not necessarily important to perform best in the study's revision task. Due to an increased focus on what is relevant for the writing tasks, students who received awareness information had more time and resources to engage in further explorations of the wiki contents.

Participants with high levels of epistemic curiosity who were in the control group without controversy awareness were inhibited in their interest to search for new stimuli in other discussions. In this current study, it could be shown that modifying wiki talk pages with visual controversy awareness information implicitly guides readers towards meaningful controversial discussions. Such controversies can be fruitful sources of information for knowledge construction processes regarding fostering beneficial socio-cognitive conflicts within learners. The study was explicitly designed for addressing learners in higher educational settings such as universities, where wikis can be deployed as course-specific closed mandatory group writing assignments. Therefore, the findings should be treated with some caution if they should be transferred to different educational contexts or even trying to implement the presented functionalities of this study into open collaborative platform such as Wikipedia or Wikiversity. To allow more generalised conclusions for these environments, quasi-experimental or even field studies with samples of the corresponding wiki audience would be valuable. Furthermore, this study was conducted as a laboratory experiment with the advantages of randomisation and control of variables, but with the disadvantages of lacking ecological validity, but not necessarily external validity in terms of population validity. Much of this study's student sample is typical for numerous student populations. Thus, it is highly likely that the fundamental effects of implicit guidance should be replicable in samples from other student populations because they do not require specific prerequisites tied to certain selection or sampling criteria. Additional guidance measures as presented can be helpful to reduce individual cognitive costs of information-seeking and coordination processes, especially for small learning groups with limited time frames to fulfil a task.

Depending on the degree of time constraints for working on a specific wiki task assignment, students might not have the minimal required time to develop a shared mental model that would render coordination as unnecessary (Kittur & Kraut, 2008). Because of such implicit guidance implementations, individual students could potentially have more free cognitive resources that they can beneficially use for collaboration. Regarding the cognitive variables that have been measured, the findings confirmed that individual variations impact the learners' seeking and selection behaviour and correspondingly have an influence on the final learning outcome. Especially the analyses on the students' Need for Cognitive Closure suggest that this variable should be considered for further research that is focused on the provision of supporting guidance in collaborative writing environments where controversies and conflicts occur.

2.3.1 Outlook

Although some promising effects could be identified with regard to learning in this experimental study and also to higher quality contributions (Heimbuch & Bodemer, 2016a), to date only cautious inferences should be drawn from the current quantitative data analyses that the provision of controversy awareness information leads to qualitatively better contributions to the wiki article and talk page threads. Therefore, more detailed content analyses of the produced knowledge artefacts are required in order to investigate if the additional support regarding the reception of this study's controversy types led to substantial differences in text production quality and elaborations of discussion replies. Furthermore, such analyses should be conducted in conjunction with the individuals' personal Need for Cognitive Closure, which was identified as a meaningful determinant to guide one's information-seeking behaviour when dealing with ambiguous contents.

The presented approach might be enriched by developing and implementing methods of automating the generation of visualisations representing controversy status information on wiki talk page discussions with the help of natural language processing (Bär, Erbs, Zesch, & Gurevych, 2011; Daxenberger, Ferschke, Gurevych, & Zesch, 2014). Collaborations in this area between computational and psychological researchers could be fruitful to draft and test cognitive group awareness tools focused on evidence-led controversies for real-world deployment opportunities, such as on Wikipedia or Wikiversity talk pages. Beyond usage scenarios in wikis, it would be interesting to investigate the potentials of controversy awareness information in other contexts where learning materials can be socially shared and discussed.

3 Representations of Authors' Community-Standing and Expertise Information on Wiki Talk Pages

Wikis are commonly used to collaboratively create user-generated content. In the creation process it is likely that controversies between authors emerge that are further discussed on an article's accompanying discussion talk page. Research gathered evidence on discussions with contradictory evidence that can be beneficial for triggering elaboration processes. It has also been shown that many wikis are not directly suited for new users to easily assess credibility and trustworthiness of provided information and identify relevant actors in a discussion. For this exploratory experimental study ($N = 61$) additional information were implemented on authors' domain-specific expertise and community-standings as social recommendation on wiki talk pages. Effects of supplemental information on the perception of controversial discussions were analysed, in consideration of relevant individual cognitive differences. The results indicate that it affects the subjective user perception of controversies, most notably if users have a high desire for rapid orientation and conclusions.

3.1 Method

3.1.1 Design and Participants

A two-group randomised experimental study was conducted to investigate the potential benefits of integrating social recommender information into wiki talk pages. The control wiki group saw the default wiki talk pages with no further modifications. The experimental group received the same wiki discussions with supplemental information on domain-specific expertise status and community-based recommendations from Wikipedia users. Dependent variables included measures on participant's perceptions of controversial wiki discussions, the correct categorisation of different controversies and time spent on each talk page. The experiment took place at the University of Duisburg-Essen (Germany) in late winter (January – February) 2015. In total, a number of $N = 65$ participants were invited to this experiment, mostly recruited from the same university. Participants were randomly assigned to either the control group (default wiki) or the experimental group (supplemented wiki). Four data sets were incomplete due to one-time local network issues, leaving the dataset with a final sample of $N = 61$ for subsequent analyses. The size of the control group was $n = 29$ and the experimental group was $n = 32$. The final sample was 78.69% female, 21.31% male and consisted of 95.08% students of various disciplines and 4.92 % were employees. The overall age ranged from 18 to 52, with an average of $M = 20.90$ ($SD = 4.70$) with one person being a relative outlier from the otherwise 18 to 29 age range ($M = 20.38$, $SD = 2.42$).

3.1.2 Materials

For this study, a total of fifteen wiki talk pages were created covering discussions with opposing points of view about diverse topics and domains. Discussion threads

were extracted from the German Wikipedia and slightly altered them to ensure that a controversy between two or more discussants was recognisable. The discussions were equally subdivided into the following three controversy categories, (1) *content-related controversies* discussing opposing viewpoints or contradictory evidence, (2) *social controversies* dealing with personal matters instead of the article's contents and (3) *formal controversies* discussing article structure rather than its content. For the discussion headings, neutral names were used such as "Discussion 2" to prevent priming towards a specific domain or controversy category. The default wiki group and the experimental wiki group had the exact same talk page contents regarding the discussions and discussants. The experimental group received modified talk page views with two additional information representations. The first additional information represented a discussant's domain-specific expertise status in the form of a badge with the discussant's rank number calculated by positive community ratings per domain (e.g. a user with high community ratings in the domain of *genetics* will also have the ranking scores transferred to the related domain of *genetic engineering* but not necessarily to discussions about *cultural anthropology*). The second additional author information called Wikes (a portmanteau of wiki and Likes) represented the author's overall user reputation or community-standing as voted by other wiki users, comparable to well-known measures in other social media platforms where individual contributions are rated (e.g. Facebook's "Like" or Google's "+1"). Since this study was meant to gain first general insights, it was decided that both types of visual information should point in the same direction, so that domain experts with the highest ranks had also the highest community-standing. Figure 3.1 shows an excerpt from a modified talk page altered with domain-specific expertise and community-standing information.

Article Talk Read Edit View history

Discussion thread 2

I have reverted the last changes again. Hyperlinking every ointless word is just plain troll behavior. Internal links should only be used for appropriate stuff like other relevant articles or technical terms. Otherwise one could also set a link to every single word.

-- **Bartretter** **177 W+** 15:50, 23. Feb. 2013 (CET)

Thank for calling me a troll, Bartretter. Plonk. Therewith you are in perfect company with the whole phalanx of psychopaths here. It is maybe the best for you that your personal talk page is restricted. But I think I will *delete* my inquiry regarding this topic.

-- **Ptorkno** **92 W+** 16:46, 23. Feb. 2013 (CET)

Have I?

-- **Bartretter** **177 W+** 16:50, 23. Feb. 2013 (CET)

Seems that you act quick like a shot.

-- **Korbleger** **97 W+** 16:55, 23. Feb. 2013 (CET)

I am proud to call myself a troll. Whoever is called a troll in here is most likely the xact opposite in reality.

-- **Happyteeth** **104 W+** 17:03, 23. Feb. 2013 (CET)

Figure 3.1. Sample page excerpt from a wiki talk page discussion with additional author information. Ranking badges reflect domain-specific expertise ratings of authors and W+ (Wikes) reflect the overall wiki author reputation regardless of the domain as rated by the community.

3.1.3 Measurements

The overall interest of this study was to investigate general effects of additionally provided author information on wiki talk pages. Therefore, several variables were measured with relations to the processing and perception of talk page discussions.

Reading times. As a process variable, reading times of the individual discussions were measured as an indicator if additional author information influences. For each discussion page, once a new page was loaded, the time difference since the previously loaded page was calculated, and these differences were averaged across pages for each participant.

Controversy categorisation. At the start of the experiment, participants were introduced to three controversy categories, 1. *Content-related*, 2. *social*, and 3. *formal* controversies. They were asked to categorise each discussion accordingly. The number of correctly categorised controversies was extracted and the numbers were averaged for further processing.

Perceived controversy resolvability. Participants had to judge how difficult they think a controversial discussion would be to resolve. The resolvability score ranged on a scale from “0 = very easy to resolve” to “5 = very difficult to resolve”. Again, an averaged score was used across all pages for further calculations.

Discussion clarity. Each participant rated the overall clarity of a discussion thread on a scale ranked from “0 = very confusing” to “5 = very clear”. For both groups, mean clarity scores were calculated over all presented discussions.

Expert agreement. Participants were asked to answer which of the discussants they agree the most with. This made it possible to compute whether participants leaned on average more towards low expertise or high expertise discussants.

Conditional effects. As a potential moderator, the participants’ Need for Cognitive Closure was measured. Thus, participants were asked to fill out the German short scale *16-NCSS* (Schlink & Walther, 2007). The questionnaire consists of 16 statements that are rated on a six-point scale about ambiguous situations, with each item ranging from fully disagree to fully agree.

User experience. To measure acceptance of the implementation of the additional author information, the *User Experience Questionnaire* (UEQ) was used (Laugwitz, Held, & Schrepp, 2008). This measurement was used to assess the usefulness of the

deployed additional information on six distinct dimensions, in comparison to the default wiki experience.

3.1.4 Study Procedure

After welcoming participants and giving them a short briefing, they were first presented with a few introductory pages to familiarise them with the general structure of wiki talk pages. This was done to ensure that every participant had at least general baseline knowledge about what talk page discussions look like and in what respect they differ from classic online forums. Subsequently, participants filled out questionnaires on socio-demographics and previous wiki experience. Before the presentation and assessment of the discussion threads began, they were given a few more instructions about the different controversy categories. Subsequently, participants were shown sample discussions from Wikipedia with additional explanation on how to categorise either controversy-related, social, or formal controversies between discussants. Additionally, the experimental group was presented with supplemental information on another introductory page that explained the meanings of badges (domain-specific user expertise) and Wikes (overall user reputation) that were implemented on the talk pages. Following this, fifteen discussions were sequentially presented without any time constraints. Each discussion thread was followed by a short assessment, in which the main dependent variables were measured (cf. Figure 3.2). At the end of the experiment, participants filled out the Need for Cognitive Closure questionnaire. Additionally, the experimental group with supplemental information filled out the user experience questionnaire.

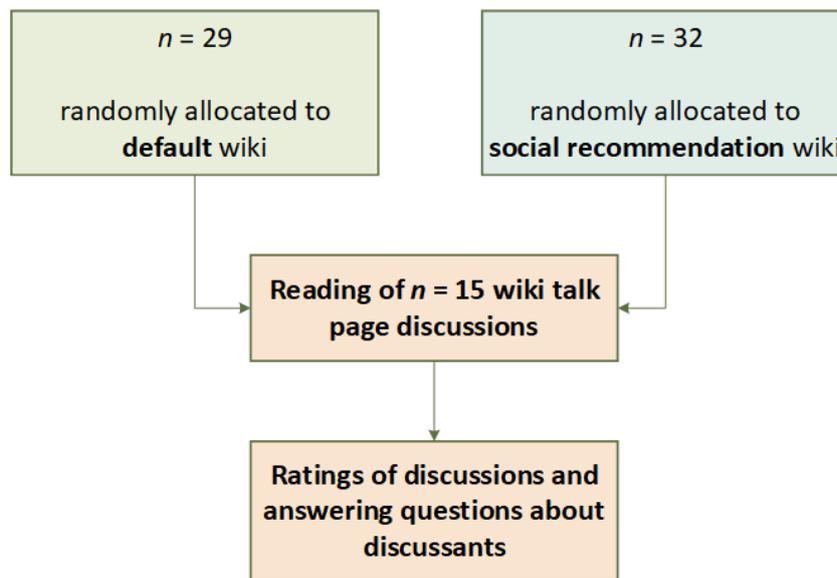


Figure 3.2. Workflow diagram visualising the overall study procedure of Experiment 2 with its central stages.

3.1.5 Specific Hypotheses to Research Questions

RQ3.1: Does the implementation of additional author information about a discussants' expertise and community-standing aid readers of wiki discussion threads in their individual assessment processes?

H3.1: It was expected that when additional author information is present, this will foster faster and better assessment of controversial discussion threads. In more detail, lower discussion reading times were expected due to visual cues for a more targeted selection of relevant actors in a discussion. Hence, It was expected that better controversy category identification will be fostered due to focused information search and selection. Furthermore, It was expected that due to this increased focus, additional author information positively influences the perception of the resolvability of a controversy and also the perception of overall clarity of discussions.

RQ3.2: How does an addition of visual author information about expertise and community-standing influence selection behaviour in wiki discussions?

H3.2: In line with current research on related social media, It was expected that users provided with additional author information are more likely to agree with an expert's line of argumentation than with a novice's argumentation.

RQ3.3: How is the perception of controversial wiki discussions influenced by the individual Need for Cognitive Closure when additional author information is present or absent?

H3.3: It was expected that when the individual Need for Cognitive Closure is low, additional information about authors is less influential. When the Need for Cognitive Closure is high, the influence is expected to be greater since these individuals predominantly prefer low level solutions and are more willing to follow visual cues without detailed assessment of content.

RQ3.4: How do users rate the additional author information implementation in terms of their user experience?

3.2 Results

Wiki expertise . Participants were asked to self-report their wiki usage frequency and experience. In this sample, passive wiki usage (i.e. browsing and reading content) was moderate with $M = 1.59$ ($SD = 0.66$) on a scale ranging from 0 (very low) to 4 (very high). The two groups did not differ in their levels of passive wiki usage, $t(58.99) = 0.71$, $p = .483$, $d = 0.18$, 95% CI [0.32, 0.68], $BF_{01} = 3.05$. Furthermore,

active wiki experience (i.e. article editing, revising articles, and participation in discussions) was low and close to zero with $M = 0.36$ ($SD = 0.47$). Between the two groups, there were no meaningful differences, $t(58.47) = 0.11$, $p = .916$, $d = 0.03$, 95% CI [-0.48, 0.53], $BF_{01} = 3.99$.

Expert agreement. As one of the study's main interests, the potential implicit guidance effects of implemented expertise ranking and community-standing information were analysed. When there was no information about discussants' expertise and reputation, as was the case in the default wiki, participants were less likely to agree with an expert's argumentation ($M = 1.69$, $SD = 1.54$) as compared to participants in the experimental wiki ($M = 3.16$, $SD = 1.90$), $t(58.27) = 3.32$, $p < .001$, $d = 0.85$, 90% CI [0.41, 1.29], $BF_{10} = 15.00$. Within the experimental wiki group there was a moderate to large effect that participants were more likely to agree with the arguments of high expertise discussants ($M_D = 1.19$, $SE_D = 0.34$), $t(31) = 3.46$, $p < .001$, $d = 0.61$, 90% CI [0.29, 0.94], $BF_{10} = 72.76$.

Reading times. As a process variable, potential group differences regarding average times spent on the wiki discussion pages were analysed when additional author information was present or absent. Participants in the default wiki spent on average $M = 1.36$ ($SD = 0.45$) minutes on a talk page compared to $M = 1.21$ ($SD = 0.32$) minutes in the experimental wiki. An independent samples test showed a moderate effect for the difference in overall talk page reading times, $t(50.54) = 1.54$, $p = .066$, $d = 0.40$, 90% CI [-0.04, 0.82], $BF_{10} = 2.93$.

Conditional effects on reading times. There was a small to moderate difference between wiki groups on single levels of the moderator variables on talk page reading times (Figure 3.3). Higher levels of an individual's Need for Cognitive Closure

led to lower reading times when they were provided with additional information, $t(57) = 1.89, p = .063, d = 0.49, 95\% \text{ CI } [-0.03, 0.99], BF_{10} = 3.41$. Furthermore, there was virtually no difference between groups when the Need for Cognitive Closure was low, $t(57) = 0.03, p = .974, d = 0.01, 95\% \text{ CI } [-0.49, 0.51], BF_{01} = 8.88$.

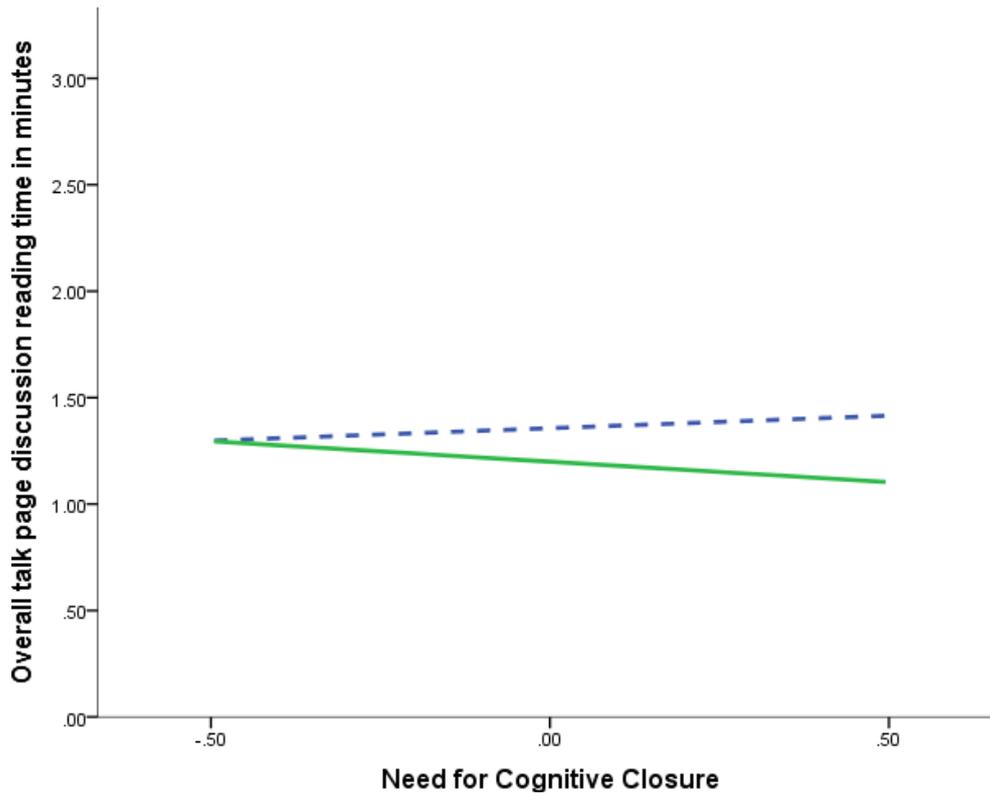


Figure 3.3. Conditional effects of the Need for Cognitive Closure on discussion reading times. Dashed lines = default wiki group, solid lines = experimental wiki group.

Discussion clarity. Participants assigned higher ratings in their subjective clarity perception of discussions when using the default wiki with an average of $M = 2.13$ ($SD = 0.86$) compared to $M = 1.69$ ($SD = 0.73$) in the experimental wiki. There was a moderate effect for a group difference on talk page clarity ratings, $t(55.28) = 2.11, p = .040, d = 0.54, 95\% \text{ CI } [0.03, 1.05], BF_{10} = 1.69$.

Conditional effects on discussion clarity. In the analyses of the Need for Cognitive Closure, there was a conditional effect of the individual Need for Cognitive Closure on the perceived clarity of a talk page, $F(3, 57) = 2.28, p = .089, \omega^2 = .06, 90\% \text{ CI } [.00, .15]$. Participants perceived controversial discussions to be less clear when additional information in form of badges and Wikes were present, but only if they had a higher Need for Cognitive Closure, $t(57) = 2.19, p = .033, \text{ Effect} = 0.70 (SE = 0.32), 95\% \text{ CI } [-0.60, -1.34]$ (Figure 3.4).

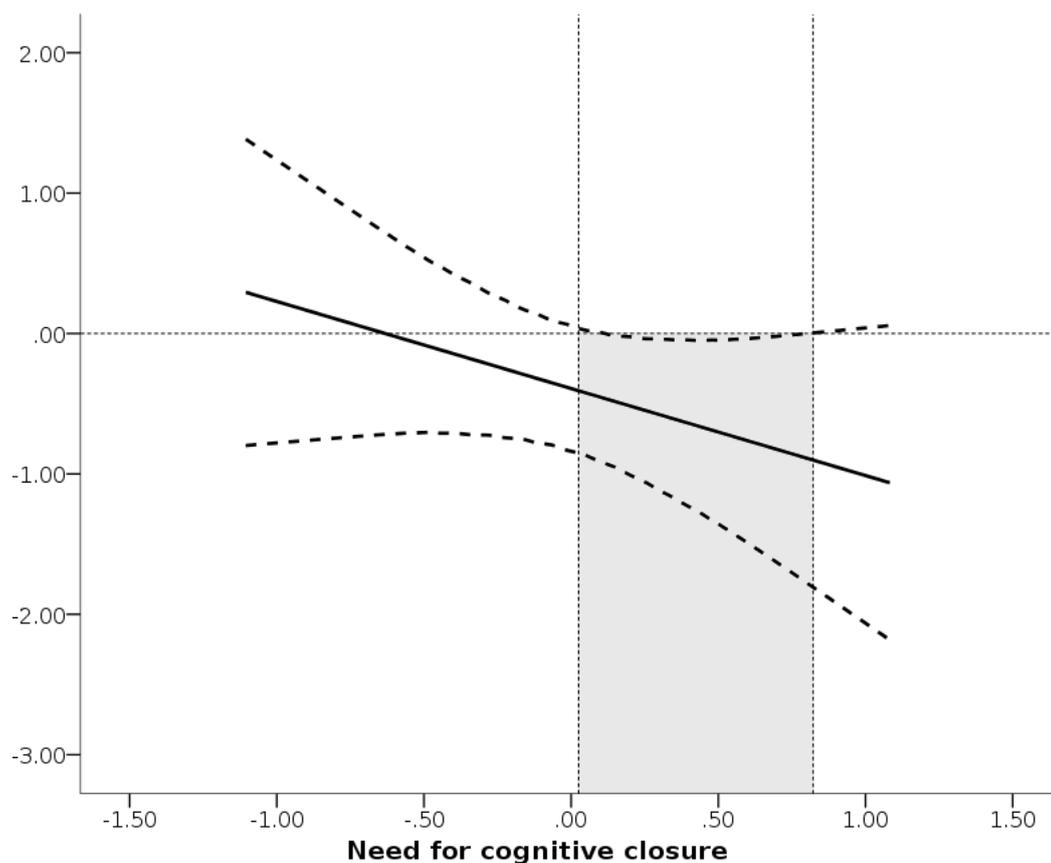


Figure 3.4. Johnson-Neyman (JN) conditional effect of social recommendations on perceived clarity of the talk page discussions moderated by the individual Need for Cognitive Closure. Dashed lines represent the upper limit (ULCI) and lower limit (LLCI) of the conditional effect's 95% confidence interval. The grey shaded area represents the JN region of significance. It is bound between the moderator's (Need for Cognitive Closure) relative values of 0.25 and 0.82 (dotted vertical lines) and below the zero line (dotted horizontal line).

Controversy categorisation. After each discussion thread, participants were asked to categorise the talk page controversy into one of three categories, resulting in $M = 10.45$ ($SD = 1.74$) correct controversy categorisation for the default wiki group and $M = 10.31$ ($SD = 1.42$) for the experimental wiki group. The group effect on correct overall controversy categorisation was negligibly small and not sufficiently supported by the observed data, $t(54.19) = 0.33$, $p = .629$, $d = 0.09$, 90% CI [-0.34, 0.51], $BF_{01} = 1.48$. More detailed analyses for the controversy categories suggest a small effect for correctly categorising content-related controversies in favour of the experimental wiki (cf. Table 3.1).

Table 3.1

Welch's t-Tests on correct controversy categorisation.

Controversy	t	df	p	d	90% CI		BF_{10}	BF_{01}
					LL	UL		
content-related	0.92	57.39	.180	0.24	0.19	0.66	1.12	0.87
social	1.64	58.31	.947	0.42	-0.01	0.85	0.46	2.16
formal	0.42	55.84	.340	-0.11	-0.53	0.32	0.92	1.09

Note. Correct controversy categorisation H_A : Default < Experimental

For the experimental group with additional information, a multiple linear regression was performed on the correct overall controversy categorisation with two predictor variables of agreement scores with either high or low ranked discussants. There was a moderate to large overall model effect of the expert agreements on correct controversy categorisation, $\omega^2 = .14$, $F(2, 29) = 3.43$, $p = .046$, $BF_{10} = 4.29$. Agreement with highly ranked discussants positively predicted a correct categorisation with moderate support by the data, $\beta = .33$, $b = 0.25$, $t(29) = 1.80$, $p = .082$, 95% CI [-0.03, 0.52], $BF_{01} = 3.93$. Conversely, agreement with low ranked discussants negatively predicted the correct categorisation with relatively strong data support, $\beta = .45$, $b = 0.40$, $t(29) = 2.47$, $p = .020$, 95% CI [-0.07, 0.72], $BF_{10} = 12.16$.

Perceived controversy resolvability. Participants of the default wiki group rated perceived difficulty of resolving a controversy minimally higher with $M = 2.15$ ($SD = 0.45$) compared to the experimental wiki with $M = 2.04$ ($SD = 0.59$), but the two groups did not differ substantially from one another, $t(57.30) = 0.85$, $p = .198$, $d = 0.22$, 90% CI [-0.21, 0.64], $BF_{10} = 1.14$. The Bayes Factor suggests that neither the null nor the alternative hypothesis model is clearly supported by the data.

Conditional effects on perceived resolvability. Regarding conditional effects of the Need for Cognitive Closure on the perceived controversy resolution difficulty, there were no meaningful group differences on any level of the moderator resulting in an inconclusive total effect model, $F(3, 57) = 0.52$, $p = .672$, $\omega^2 = .02$, 90% CI [.00, .08], $BF_{10} = 1.56$. Similar to the previous unconditional analysis, the Bayes Factor suggests that neither the null nor the alternative hypothesis model is clearly supported by the data.

User experience. Analysing the six dimensions of the user experience questionnaire UEQ showed that participants in the study's experimental wiki group rated the additional information of expert rankings and community-standing on wiki talk pages as perspicuous, $M = 1.12$, 95% CI [0.69, 1.55]. To a lesser extent, but still relevant, participants rated the visualisations as slightly efficient, $M = 0.38$, 95% CI [0.00, 0.77]. On the remaining dimensions of attractiveness ($M = 0.09$, 95% CI [-0.24, 0.42]), dependability ($M = 0.18$, 95% CI [-0.15, 0.51]), stimulation ($M = 0.12$, 95% CI [-0.29, 0.53]) and novelty ($M = -0.22$, 95% CI [-0.59, 0.15]), the ratings were inconclusive with confidence intervals of the mean dimension ratings spanning from negative to positive values (cf. Figure 3.5).

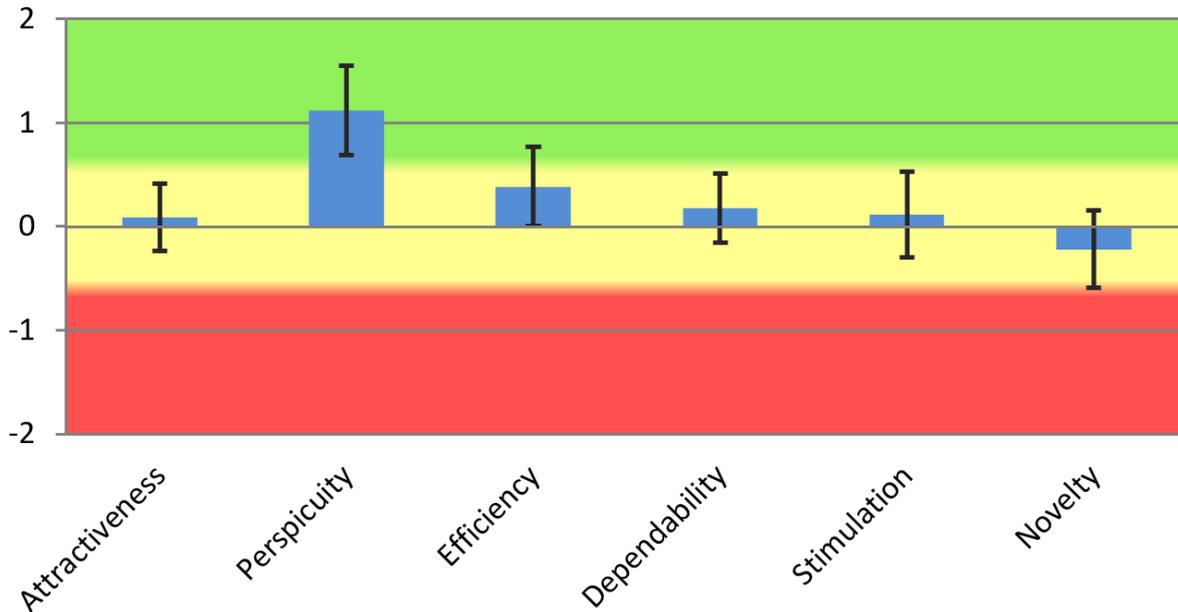


Figure 3.5. Scale results of the six dimensions of the user experience questionnaire UEQ for the additional information (ranking badges and Wikes) as used in the experiment's modified wiki.

3.3 Discussion

Using social knowledge construction platforms such as Wikipedia as a starting resource for finding high-quality content, it can be necessary for an individual user to assess quality in terms of credibility and expertise of the source. Due to the quantity of information and the lack of effective default filtering mechanisms, users can feel lost in the system. Thus, additional guidance seems to be an obvious necessity to aid in finding and selecting what is relevant and who the most important actors are in terms of credibility and expertise in the specific domain(s) of interest. Previous research on external representations promoting implicit guidance on wiki talk pages has gathered evidence for their effectiveness and efficiency in terms of individual learning and knowledge construction artefacts (Heimbuch & Bodemer, 2017), but without addressing how credible an author or the provided content was.

With this presented study it was aimed to gain some insight into credibility assessment processes and the subjective perception of controversies and conflicts occurring in wiki discussions. In analysing the present data, it has been found that when users had additional information about the community-standing of a discussant in form of rank badges and general community agreement, they were more likely to agree with an expert's line of argument. The computed Bayes Factors of about 15 and 73 for the main analyses represent very strong support for the potential of the implemented expertise rank and community-standing information to guide users on wiki talk pages. Further support for benefits of the additional information was obtained when the data on controversy categorisation was analysed, where users performed better in correctly categorising a controversy if they followed the argumentation of the highest ranked discussion expert, and with the adverse effect on controversy categorisation if they followed the lower ranks. This adds to previous research on high and low status sources that trigger different modes of social comparison heuristics when dealing with socio-cognitive conflicts (Baumeister, Englemann, & Hesse, 2016). The expertise and community-standing information aided participants in categorising controversies especially if the controversy was content-related. The visualisations also altered their perception of the whole discussion by causing them to consider controversies to be easier to resolve when one of the discussion's experts had been asked for assistance. Regarding the help-seeking process, this could have the potential to massively ease the decision process of whom to solicit help from if one gets stuck in lengthy discussions with contradictory evidence and opposing points of view (Makara & Karabenick, 2013). Furthermore, users in the experimental wiki group with additional information spent less time with the presented discussions. Thus, it is possible that they needed less time for orientation in the viewed discussion threads.

Reducing the time that individuals need to get a first overview and focus on relevant actors of a discussion is desirable since their attention is directed towards the parts of a discussion that are most valuable to them (Lops, Gemmis, & Semeraro, 2011; Marinho et al., 2011). It can be assumed that the expertise status ranks and community rating information helped participants to focus more selectively on discussants who might be the most significant actors in a discussion. The more accurately wiki users recognise controversies as constructive and relevant, the more likely it becomes for meaningful socio-cognitive conflicts to be induced when different or prior knowledge about the discussed subject matter exists (Chi, 2008; Mugny et al., 1995). Because of the prevalence of opposing evidence and opinions in discussions, there was also considerable interest in analysing a cognitive construct related to ambiguity (i.e. Need for Cognitive Closure) and the role that it plays. The data suggest that participants with a low Need for Cognitive Closure do not differ in their judgements about the perceived clarity of a discussion regardless of additional information. In contrast, when one has a higher one Need for Cognitive Closure, additional information of expertise alters the participants' judgements towards a higher perceived clarity of the discussions. Moreover, users with high levels of Need for Cognitive Closure spent less time reading discussion threads when the expertise and community ratings were present. This in accordance with previous research that suggested that high levels of Need for Cognitive Closure led to the urge for quick but plausible answers rather than the best possible solutions to a problem or state of ambiguity (Webster & Kruglanski, 1994). Both types of supplemental information seem to support the judgement of credibility processes for high closure individuals, but they show virtually no effect for low closure participants.

Although findings regarding Need for Cognitive Closure are quite noisy due to the likelihood of high sampling error of the medium-sized sample, it provides valuable insights in further information on the effects of source cues and social recommendation in Web 2.0 environments, going beyond previous research on credibility (Winter & Krämer, 2014). Finally, regarding the user experience of the supplemental information, the representations of expertise rankings and community-standing were not rated as novel by the study participants. This is not much of a surprise given that participants were recruited mostly out of cohort prolific to the use of Facebook, Twitter, and similar participatory websites. But users generally valued the presented information as easily understandable in terms of perspicuity and even regarded them as slightly efficient. Both information of domain-specific expertise and community-standing seem to provide quick and unobtrusive tacit guidance to categorise a controversial discussion and to identify potentially helpful experts in a discussion, which one could use to seek further advice or be nominated as a potential mediator to resolve controversies in a certain domain. In conclusion, it seems that discussions appear clearer and more comprehensible when additional information reflecting expertise status of wiki authors is present on discussions, thus making them more accessible to a wider audience. Nevertheless, from this exploratory study it remains unclear which of the information had what strength of guiding effects. Further studies should aim to disentangle and quantify these effects for both domain-specific expertise and community-standing information. The results regarding the assessed influencing variables suggest that further investigations of individual differences should be considered in future larger scale replications. Furthermore, discussions such as the ones used in this experiment have the potential to induce socio-cognitive conflicts between the information in the wiki and an individual's knowledge and also between diverging knowledge of multiple contributors

(Cress & Kimmerle, 2008; Heimbuch & Bodemer, 2017). Such conflicts can trigger elaboration and learning processes which result in more diverse individual knowledge by opening new perspectives for readers and contributors, and thus should be examined in further research on wikis and similar environments.

4 Investigating Collaboration Scripts for Learning with Wiki-Based Environments

Knowledge construction assignments with wikis can be found in various educational settings, but the environment is not inevitably suited to facilitate learning and thus requires additional guidance. In this study, the effects of two collaboration scripts with different objectives were investigated during a two-week period of knowledge construction with wikis as a supplement to lectures about descriptive statistics. One script that is derived from a workflow as suggested by Wikipedia promoted a high frequency of individual article edits without further coordination. In contrast to this approach, an alternative script encouraged participants to discuss any planned changes upfront. Results indicate that a discussion collaboration script proposal encouraged students to participate in the whole script process, while in Wikipedia's script proposal only the first step of article editing was executed. These edits were generally of slightly lower quality. Learning success was not directly affected by the scripts, though the data suggests small effects in favour of the discussion script proposal.

4.1 General Method of Experiments 3 & 4

Both experiments integrated BRD and DDR script representations in either a self-developed wiki-mimicking learning environment in Experiment 3 and a Wikipedia-inspired *DokuWiki* in Experiment 4. Prior to working on the collaborative wiki writing tasks, students were introduced to the respective operational sequences of the BRD and DDR script. In both experimental studies participants were rewarded with a certificate of attendance needed during their studies. Students were randomly assigned to one of the two collaboration script groups in both experiments, as complementary learning partners in groups of two in the first study and individually in the second study. The primary task in both experiments was to collaboratively edit existing wiki contents that have either been generated by the experimenters for Experiment 3 or utilising student-generated wiki contents from a previous study as a foundation for Experiment 4. Despite all differences in settings and setups of both studies, participants in both experiments were asked to fill out the 16-NCCS (Schlink & Walther, 2007), a short questionnaire on the individual need for cognitive closure. The questionnaire consists of 16 statements to be rated on a six-point scale about how one deals with ambiguity, each item ranging from “fully disagree” (1) to “fully agree” (6). Additionally, in Experiment 4 further measurements of influencing variables have been conducted that were not previously addressed in Experiment 3 (cf. 4.2.3 and 4.4.3).

4.1.1 Specific Hypotheses to Research Questions

The interested research questions, as previously discussed in 1.8, and the tested hypotheses of both Experiments 3 and 4 on collaboration scripts are visually represented in the following Figure 4.1.

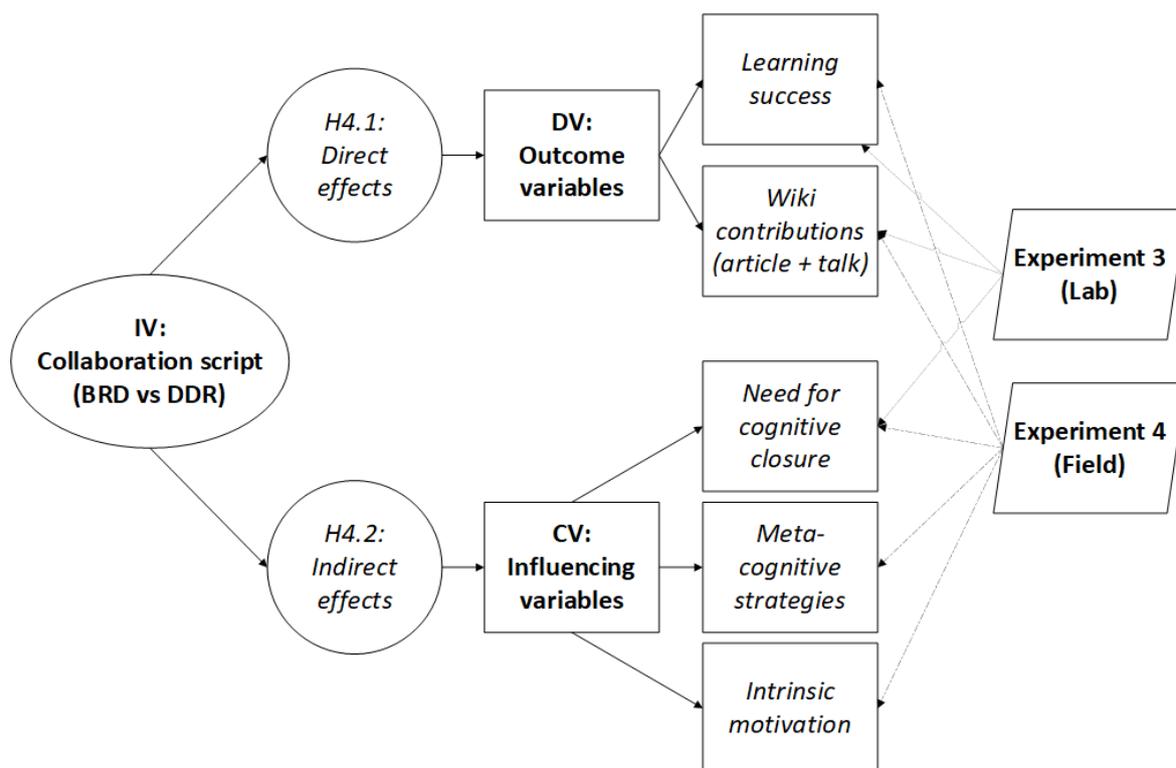


Figure 4.1. Hypothesis model representing which of the experiments (3 and 4) addressed what effects, dependent variables and influencing variables.

RQ4.1: Does a DDR script have a greater positive impact on learning and the wiki article quality than Wikipedia’s proposed BRD script does?

H4.1a: A DDR script leads to higher knowledge than a BRD script.

H4.1b: A DDR script leads to a higher participation than a BRD script.

H4.1c: A DDR script leads to fewer article edits than a BRD script.

H4.1d: A DDR script leads to higher quality articles than a BRD script.

RQ4.2: To what extent do the individual cognitive variables influence the effects on knowledge construction and the frequency of wiki activities in the talk page discussion forums?

H4.2a: The script's effects on the learning outcome are affected by the learners' need for cognitive closure, their metacognitive strategies and their intrinsic motivation.

H4.2b: The script's effects on the contributions are affected by the learners' need for cognitive closure, their metacognitive strategies and their intrinsic motivation

4.2 Method of Experiment 3

4.2.1 Design and Participants

To test the hypotheses, the first script study was conducted as a two-group between-subjects experiment with participants working in dyads. At the start of the laboratory experiment (t1), the sample consisted of $N = 28$ university students (14 females, 14 males) with a mean age of $M = 23.29$ ($SD = 5.72$). All experiments were conducted in a research laboratory where participants and were seated at computers separated by privacy blinds. They were randomly assigned to a BRD or DDR script group and received further instructions on the computer. Participants rated their topic-specific interest as medium to high (broader topic: $M = 3.50$, $SD = 1.11$; specific topic: $M = 3.14$, $SD = 1.08$) and their self-assessed prior knowledge about the topic ranged between low and moderate (broader topic: $M = 2.00$, $SD = 1.02$; specific topic: $M = 0.25$, $SD = 0.80$) (cf. 4.2.3 for information on the scales used). Between groups, there were negligibly small to no differences in topic-specific interest (broader topic: $U = 62.00$, $z = -1.71$, $p = .104$, $r = -.32$, 95% CI [-.62, .06]; specific topic: $U = 73.00$, $z = -1.12$, $p = .265$, $r = -.23$, 95% CI [-.55, .28]) or in prior knowledge (broader topic: $U = 72.00$, $z = -1.27$, $p = .246$, $r = -.24$, 95% CI [-.56, .15]; specific topic: $U = 91.50$, $z = -0.56$, $p = .769$, $r = -.11$, 95% CI [-.46, .28]). Two weeks after the

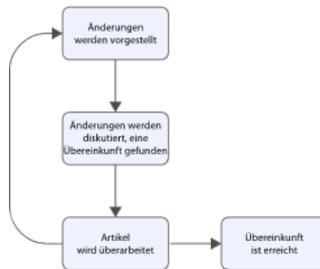
lab study (t2), participants were asked to take part in an online post-test about the study's contents. $N = 22$ students (12 females, 10 males) accepted to participate in this post-test.

4.2.2 Procedure and Materials

After being welcomed and declaring their consent, participants were placed in front of a computer where they received further instructions to start the experiment. First, they were asked to provide demographic information (e.g. gender and age) and to rate their interest in and prior knowledge about the study's controversial topic. The broader topic was pirates in general and more specifically a historical personality, namely the alleged pirate Captain William Kidd. Participants then read a wiki article about the Scottish-American buccaneer William Kidd and supplementary material that was presented to the participants individually. In each learning dyad, one participant received (A) evidence that supported the theses that William Kidd was a pirate and his trial was justified and the other received (B) evidence against those theses. In a mandatory tutorial, they were then introduced to the phases and workflows of either the BRD or the DDR script. Detailed instructions were given as texts and accompanied by graphical representations of the most relevant phases of the scripts. Following this, representations of the collaboration scripts as work flow diagrams depicting the most relevant phases were permanently visible in the learning environment (cf. Figure 4.2). Learning partners were invited into a learning environment with the collaborative online open source editor *Etherpad*. Prior to the experiment's collaboration phase, participants had the opportunity to familiarise themselves with Etherpad and its functions in a prepared tutorial. The study's main task in both groups was to edit a wiki article about William Kidd with the help of the supplementary materials each participant received.

Artikel: „William Kidd“

Bitte bearbeiten Sie den Artikel anhand der ausgehändigten Zusatzinformationen. Halten Sie sich dabei möglichst nah an den „Talk first, Consensus, Revise“-Zyklus.



The screenshot shows a text editor interface with a toolbar at the top containing icons for bold, italic, underline, strikethrough, bulleted list, numbered list, link, unlink, undo, redo, and a search icon. The text area contains the following text:

1 William Kidd (* 1645 in Greenock; † 23. Mai 1701 in London) war
 2 ein schottisch-amerikanischer Freibeuter, der später aber als
 3 Pirat zum Tode durch den Strick verurteilt wurde.
 4
 5 **Leben**
 6 Kidd wurde in Schottland geboren, emigrierte später nach
 7 Amerika und ließ sich in New York City nieder. Dort heiratete er
 8 Sarah Bradley Cox Oort. Sie hatten zwei Töchter miteinander:
 9 Elizabeth und Sarah Kidd. Die Heirat brachte Kidd ein
 10 beträchtliches Vermögen, und bevor er Pirat wurde, war er ein
 11 respektierter Kaufmann.
 12 Während einer Fahrt nach England wurde Kidd eine
 13 Freibeuterlizenz angeboten, mit der er französische Schiffe
 14 kapern durfte. Vier Fünftel der Kosten wurden durch adlige
 15 Herren bezahlt, die mit zu den mächtigsten des Königreiches
 16 gehörten; der Earl of Orford, der Baron von Romney, der
 17 Herzog von Shrewsbury und Sir John Somers. Diese handelten
 18 im Namen der British East India Company. Kidd und ein
 19 Bekannter, Oberst Robert Livingston, bezahlten den Rest. Kidd
 20 musste sein Schiff „Antigua“ verkaufen, um die Mittel
 21 aufzutreiben.
 22 Das neue Schiff, der Dreimaster „Adventure Galley“, war gut
 23 geeignet für die Jagd auf Piraten; es hatte 34 Kanonen und 150
 24 Mann Besatzung. Kidds Unternehmen war jedoch kein Erfolg
 25 vergönnt. Er hatte große Schwierigkeiten, die Kosten zu decken
 26 und geriet unter Druck seiner Männer, er solle jedes Schiff
 27 angreifen, das nicht englisch war. Rechtlich durfte er jedoch nur
 28 französische Schiffe und Piraten angreifen.
 29 Am 30. Oktober 1697 kam es zu einem Streit mit dem
 30 Geschützmeister William Mearns. Dieser Streit endete in einem

Figure 4.2. Wiki-mimicking learning environment with collaboration script representation in Experiment 3.

In the collaboration phase a basic article about William Kidd was preloaded into Etherpad. This article was structured in a way to mimic a standard Wikipedia article and had a total length of 810 words. Because the lab experiment was constrained in time, Etherpad’s built-in chat functionality was activated by default and worked as a substitute for the corresponding wiki article talk page. After the collaborative editing task, participants were again asked to answer questions about their interest in the study’s topic and the likelihood that they will search for more information after the experiment. This was followed by a short questionnaire on their individual need for cognitive closure. As an immediate test of what participants had learned about the topic, a knowledge test with multiple choice and open-ended questions

was deployed at the end of the session (t1). Two weeks later participants were invited to participate in a second knowledge test (t2) with a different but comparable set of multiple choice and open-ended questions. Figure 4.3 shows the overall study procedure.

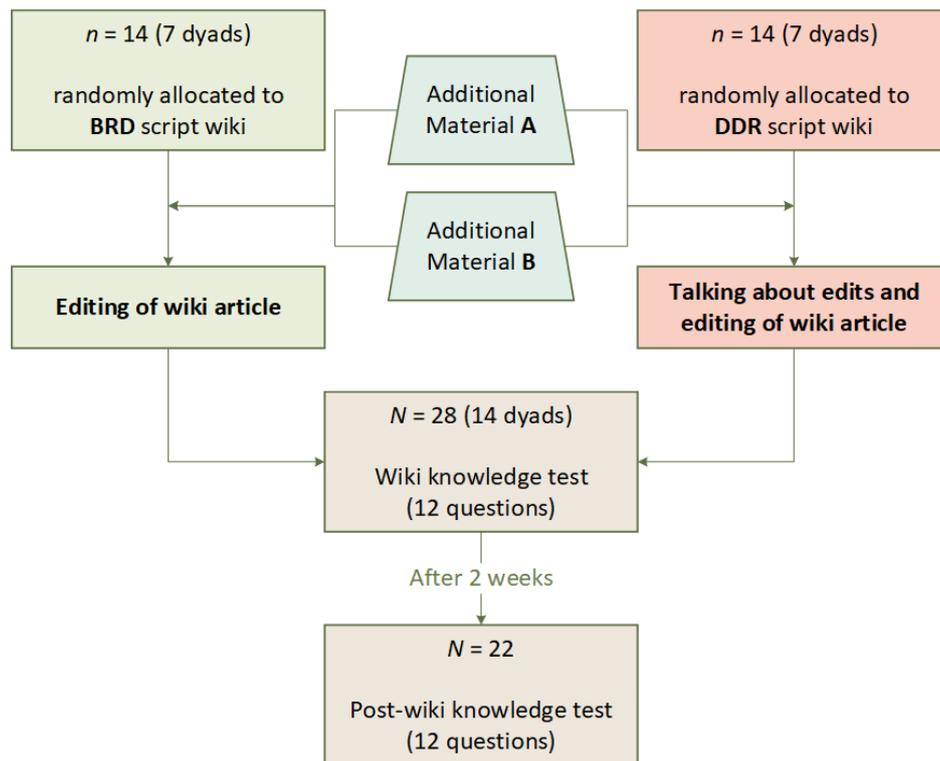


Figure 4.3. Workflow diagram visualising the overall study procedure of Experiment 3 with its central stages.

4.2.3 Variables and Measurements

The main independent variable was the random allocation to either of two experimental wiki-like environments proposing different collaboration scripts. While participants of one group were assigned to the collaborative editing workflow proposed by Wikipedia (BRD), the other group was assigned to the script that was designed with a stronger focus on upfront discussion and deliberation processes (DDR).

The learning outcomes were measured at two points in time, immediately at the end of the experiment (t1) and two weeks later with an invitation-only online questionnaire (t2). This leads to a 2 x 2 mixed factorial design with one between- and one within-subjects factor with individuals as unit of analysis, $ICC = -.07$, $F(13, 14) = 0.87$, $p = .598$. In addition to socio-demographic variables (e.g. gender and age), topic-specific interest and prior knowledge about the study's topic were assessed. Participants were asked to provide self-ratings on six-point scales ranging from "not interesting at all" / "no prior knowledge" (0) to "very interesting" / "high prior knowledge" (5). To measure the impact either collaboration script had on the knowledge construction processes and learning outcomes, wiki contribution metrics were recorded by the wiki environment and two knowledge tests were deployed. Contributions to the wiki were analysed in terms of article length, discussion length and the number of ideas from the supplementary materials that were incorporated into the original article. Measures of article and discussion length functioned as quality indicators, but also as manipulation checks to infer if participants collaborated in accordance with the respective script. Between the collaborative writing task and the first assessment of learning, the personal need for cognitive closure was measured with the German short scale 16-NCCS. The two knowledge tests to measure the learning outcomes each comprised ten multiple choice questions with four alternative answers and two open-ended questions. In each of the multiple-choice tests, a maximum score of 17 correct answers could be achieved. In the open-ended questions participants were asked if they think that (1) Captain William Kidd was a real pirate and (2) if his conviction was justified. These questions were used to assess whether a participant provided only arguments presented in the individual supplementary material, or if they gave more differentiated statements by integrating the evidence from their partner's material.

4.3 Results of Experiment 3

4.3.1 Direct Group Comparisons

Regarding the participants' contributions, the article lengths was analysed at the end of the collaboration, the integration of new evidence from the supplementary materials as well as individual chat logs. The chat log was also meant to serve as a manipulation check to infer from the intensity of the discussions if dyads have followed their respective collaboration script. One dyad was excluded from the chat analyses because of inappropriate behaviour. Due to the very small sample non-parametric Mann-Whitney U-tests were performed to analyse relevant differences between the script groups in the above-mentioned contribution variables.

Article revisions. For the whole sample, the article revisions resulted in a mean length of $M = 1210.50$ ($SD = 127.27$), in comparison to the original article's length of exactly 810 words. Comparing revised article lengths of the BRD script group ($M = 1202.57$, $SD = 99.43$) with the DDR group ($M = 1218.43$, $SD = 153.69$) the data did not suggest a meaningful difference between the script groups, $U = 96.00$, $p = .945$, $r = .02$, 95% CI $[-.39, .43]$, $BF_{01} = 1.06$. Participants in the BRD group added marginally more ($M = 9.43$, $SD = 2.15$) additional evidence into the article than the DDR group ($M = 8.43$, $SD = 2.82$), but an analysis of the contribution frequency difference did not indicate a meaningful effect, $U = 120.00$, $p = .318$, $r = .22$, 95% CI $[-.20, .58]$, $BF_{01} = 1.60$.

Discussions. Regarding the usage of the chat as a substitute for the article's talk page, there was a substantial difference between the BRD script group ($M = 133.64$, $SD = 94.01$) and the DDR group ($M = 283.08$, $SD = 186.35$) with a moderate to large effect, $U = 58.00$, $p = .035$, $r = .41$, 90% CI $[-.07, .66]$, $BF_{10} = 2.45$.

Learning outcomes. To gain insights about the effects on learning success, the participants' knowledge was assessed the first time during the experiment after the collaborative writing phase (t1). To measure potential differences in the consolidation of knowledge, a second measurement of the learning outcomes was conducted two weeks after the original experiment (t2). Overall, participants collaborating with the BRD script achieved a moderately lower score of $M = 23.67$ ($SD = 2.06$) compared to the average score of $M = 25.54$ ($SD = 3.89$) in the DDR script group, $U = 32.50$, $p = .043$, $r = .44$, 90% CI [-.06, .72], $BF_{10} = 1.43$. In greater detail, in the first knowledge test at t1 learners in the BRD scripting group tended to score lower ($M = 12.33$, $SD = 1.12$) than those in the DDR group ($M = 13.46$, $SD = 2.96$), $U = 66.50$, $p = .074$, $r = .32$, 95% CI [-.03, .60], $BF_{01} = 1.17$. In the second test at t2 the group difference was reduced, but still favoured the DDR script ($M = 12.08$, $SD = 1.61$) over the BRD script ($M = 11.33$, $SD = 1.00$), $U = 40.50$, $z = 1.26$, $p = .111$, $r = .31$, 95% CI [-.10, .63], $BF_{10} = 1.37$. In direct comparison, learners working on an article according to the DDR script achieved higher test scores immediately after the collaboration and this difference in favour of the DDR script remained relatively stable after two weeks.

Perspective-taking. Regarding the open-ended questions, they were categorised whether a participant answered these questions by means of one's own supplementary material or by integrating evidence from the learning partner's supplementary material. For the first question regarding Captain William Kidd's alleged pirate status groups differed largely in their answering behaviour at t1, $\chi^2(1, N = 28) = 7.34$, $p = .007$, $BF_{10} = 15.19$ and remained relatively stable at t2, $\chi^2(1, N = 22) = 2.20$, $p = .138$, $BF_{10} = 1.32$. The odds of incorporating evidence from the partner's learning material into one's own answer was 10.80 times higher at t1 and 4.08 times higher at t2 if students collaborated with the DDR script (cf. Figure 4.4). In conclusion, when

measured closely after the collaboration, learners who had worked with the DDR script were much more likely to provide complex answers using arguments from their own as well as their partner's additional materials. This effect was reduced after two weeks, but the likelihood of providing an answer that integrates two opposing viewpoints was still higher for learners who previously collaborated within the DDR script group.

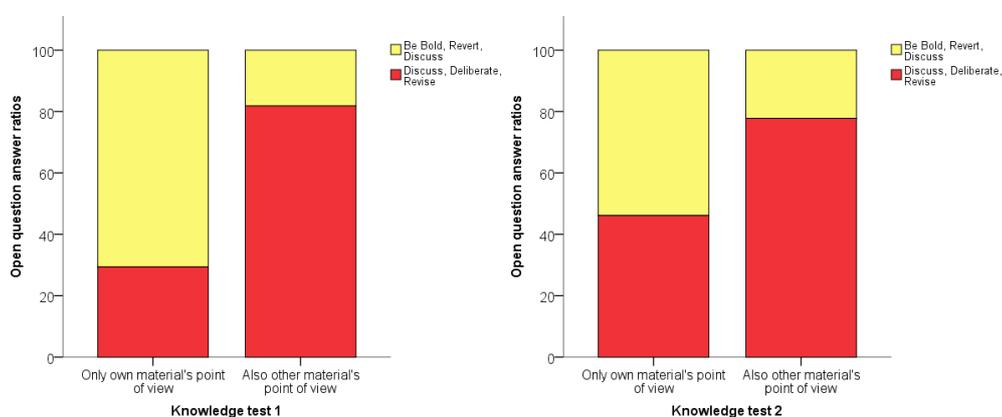


Figure 4.4. Incorporation of viewpoints of one's own learning material and the learning partner's material at t1 (left) vs t2 (right).

Regarding the second open-ended question about William Kidd's conviction, there were absolutely no relevant differences in the answering behaviour at t1 and t2, $\chi^2(1, N = 27) = 0.07, p = .785, BF_{01} = 2.29$ and $\chi^2(1, N = 20) = 0.59, p = .444, BF_{01} = 1.59$, respectively. The odds of incorporating evidence from the partner's learning material into one's own answer was 0.80 times higher at t1 and 2.14 times higher at t2 if they collaborated using a DDR script. The differences between answering behaviour on the two open-ended questions can be explained by two factors. First, the question about the conviction might have been more complex. Second, the number of arguments provided in the supplemental materials was lower for the conviction justification than for the materials concerning the pirate status discussion.

4.3.2 Influencing Variables

Need for cognitive closure. To explore the potential effects of learners' need for cognitive closure (NCC) on the contribution behaviours and the learning outcome, Spearman's rank-correlations ρ and scatterplots were used (cf. Figure 4.5). For the BRD group there was a negative correlation between the total number of newly added evidence and the NCC, $\rho(12) = -.31, p = .139, 90\% \text{ CI } [-.67, .17], BF_{01} = 1.02$. In contrast, for the DDR group there was a positive correlation between the NCC and the amount of added evidence, $\rho(12) = .29, p = .155, 90\% \text{ CI } [-.19, .66], BF_{01} = 1.17$. A similar trend was identified when the correlations between article length and the NCC were analysed, resulting in a negative correlation for the BRD group and a positive correlation for the DDR group, $\rho(12) = -.34, p = .119, 90\% \text{ CI } [-.69, .14], BF_{01} = 1.32$ and respectively, $\rho(12) = .36, p = .107, 90\% \text{ CI } [-.12, .70], BF_{01} = 1.05$. Regarding chat frequency, there was a small to moderate correlation for the NCC in the BRD group, $\rho(12) = .25, p = .191, 90\% \text{ CI } [-.23, .64], BF_{01} = 1.63$ and conversely a small to moderate negative correlation in the DDR group, $\rho(10) = -.17, p = .286, 90\% \text{ CI } [-.61, .36], BF_{01} = 1.41$. The aggregate learning outcome of both knowledge tests was positively associated with the NCC for the BRD group, $\rho(7) = .26, p = .250, 90\% \text{ CI } [-.38, .73], BF_{01} = 2.16$ and slightly negatively correlated with the NCC for the DDR group, $\rho(11) = -.14, p = .326, 90\% \text{ CI } [-.58, .36], BF_{01} = 2.52$. The Bayes Factors for all analyses on the Need for Cognitive Closure give at best anecdotal support for the null hypotheses of no differences. It is evident that the underlying data of this fairly small study sample is not sensitive enough to draw any clear statistical conclusions regarding the Need for Cognitive Closure.

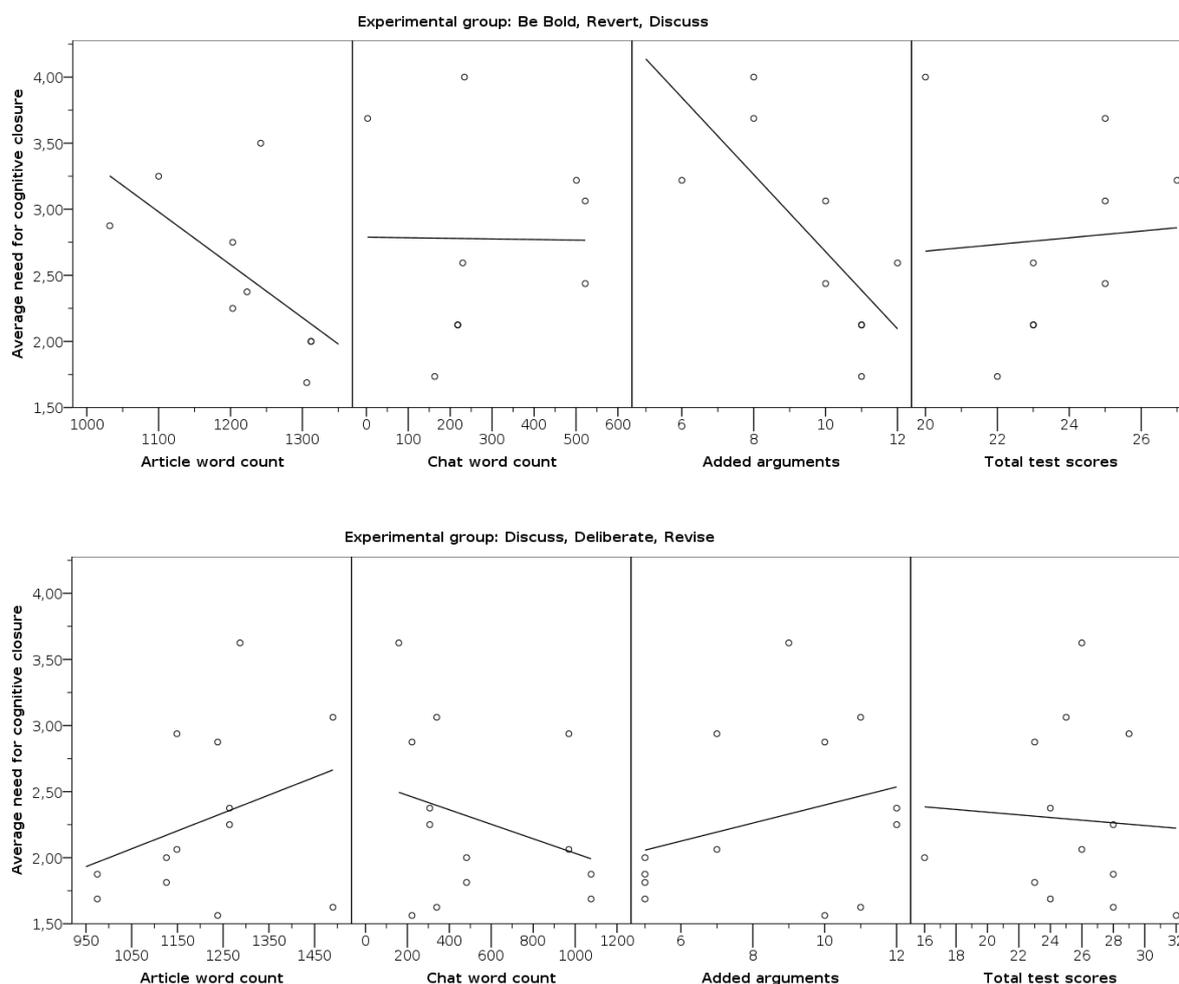


Figure 4.5. Matrices of need for cognitive closure scatterplots for Be Bold, Revert, Discuss script (top panel) and Discuss, Deliberate, Revise script (bottom panel).

4.3.3 Summary of Main Findings

The central result of Experiment 3 is that when dyads followed their respective collaboration script, the resulting outcomes and processes differed substantially. Students in the DDR script group were encouraged to do more coordination and exchange more information before proceeding to edit the original article. The articles themselves did not differ between script groups, neither in text length nor in the amount of newly added arguments. Regarding the learning outcomes and acquisition of the partner's arguments, the students showed some meaningful differences

depending on their script. Students in the DDR group scored slightly higher on average in the knowledge tests, but more impressively they were much more likely to incorporate additional information and arguments from their learning partner's material. The correlations of the need for cognitive closure with the dependent variables did not show any clear patterns, which is likely due to the small sample and thus a relatively high sampling variability. Overall, put cautiously, it seemed that the DDR script encouraged students who are high on the need for cognitive closure scale to be more engaged in the knowledge construction process by writing slightly longer articles and adding more arguments to the text.

4.4 Method of Experiment 4

4.4.1 Participants and Design

Ninety students in a course on Inferential Statistics for Applied Cognitive and Media Science were invited to participate in Experiment 4 and completed the pre-test. Due to attrition, $N = 69$ of them completed all stages of the experiment. Participants were aged 18 to 27 years ($M = 20.59$, $SD = 2.02$) and consisted of 56 (81.16%) female and 23 (18.84%) male students. The study's main collaboration phase was conducted as a randomised experiment in a natural wiki setting, so that students could contribute to the wikis from any place at any time. The average pre-test score in the DDR group was $M = 12.91$ ($SD = 1.36$); students in the BRD group achieved $M = 13.39$ ($SD = 1.71$) points on average. The group difference regarding prior knowledge as reflected in the pre-test score was small and not substantial, $t(58.96) = 1.26$, $p = .211$, $d = 0.31$, 95% CI [-0.17, 0.78], $BF_{10} = 1.95$.

4.4.2 Materials and Procedure

The basis for the wiki materials were created in a previous study, conducted the previous semester by the same student cohort (Mock, 2017). In that study, students worked in small groups of two to four people. They initiated ten wiki pages and wrote nine articles about statistical topics, one topic was left with a blank article template (Table 4.1). The resulting wiki articles were used as the original articles in the current experiment.

Table 4.1

Titles and lengths of the original wiki articles that were created in the previous semester.

Topic titles	Word count
Measures of dispersion	758
Degrees of freedom	181
Hypotheses	284
Confidence intervals	407
Measures of central tendency	0
Levels of measurement	767
Significance testing	521
<i>t</i> -Test	354
Distributions & <i>z</i> -Transformation	225
<i>z</i> -Test	388

The main task for students in this study was to review and revise these articles during a specified period. The study was advertised to the students as an optional supplement to the tutorials of their statistics course. Participants were guided to expand upon and improve the existing content qualitatively and quantitatively. In contrast to the previous lab Experiment 3, participants did not receive any additional learning materials. Instead, they were instructed to use their newly acquired

and more detailed knowledge gained from the follow-up statistics course in the second semester. It was aimed to stimulate subject-related as well as formal conflicts between students' previous knowledge and to gain new insights into statistical topics. The study was divided into three phases: (1) a knowledge pre-test about descriptive statistics and additional questionnaires including relevant socio-demographics, (2) a two-week time frame for collaborative wiki revisions and (3) a knowledge post-test about the same statistical topics (cf. Figure 4.6).

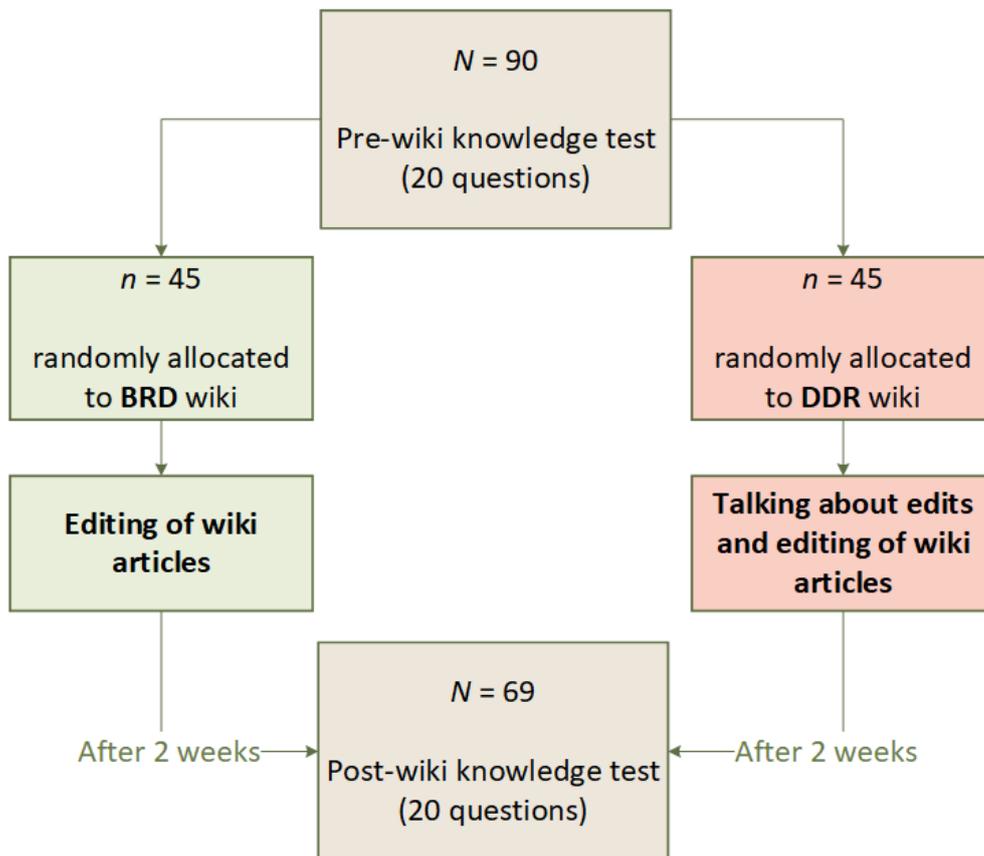


Figure 4.6. Workflow diagram visualising the overall study procedure of Experiment 4 with its central stages.

Both groups were provided with the same wiki article contents. To conduct this experiment with larger student groups, two DokuWiki instances served as experimental wikis in which either the BRD or the DDR script was included. Students were first led to a start page where they were provided with detailed instructions about the respective work flows of the collaboration scripts. As permanently visible reminders of the most relevant steps, graphical script representations appeared in the header of each wiki page. In addition, the scripts' main stages were also included as textual representations in the wikis' navigation panes (cf. Figure 4.7).

The screenshot shows a DokuWiki interface. At the top, there are navigation tabs: 'Artikel', 'Diskussion', 'Lesen', 'Bearbeiten', and 'Ältere Versionen'. A search bar is located to the right of these tabs. Below the tabs, a small diagram is visible, showing a flowchart with several boxes and arrows. The main content area contains the following text:

Vielen Dank, dass du weiterhin an meiner Studie teilnehmen willst!

Im Rahmen meiner Arbeit beschäftige ich mich mit Unterstützungsmaßnahmen Wiki-basierter Lernumgebungen.

Diese Studie wird in Zusammenhang mit meiner Bachelorarbeit am Lehrstuhl Psychologische Forschungsmethoden durchgeführt. Du befindest dich nun bei Teil 2 meiner Studie. Die Teilnahme wird noch ca. 3.5 Stunden dauern und insgesamt werden 4 Versuchspersonenstunden vergeben. Die im Laufe der Studie gemachten Angaben werden ausschließlich für wissenschaftliche Zwecke verwendet und nicht an Dritte weitergegeben. Melde dich bitte mit deinen per E-Mail erhaltenen Log-In-Daten an.

Ziel dieser Studie ist es zu erfahren, wie Personen mit Wikis gemeinsam lernen und wie man diesen Lernprozess bestmöglich unterstützen kann. Dazu findest du in diesem Wiki nun zehn Themen, die du bereits aus der Deskriptiven Statistik kennst. Diese können unvollständig oder sogar fehlerhaft sein. Ziel ist es nun von deinen Kommilitonen und dir, dieses bereits bestehende Wiki gemeinsam zu erweitern und zu optimieren. Du kannst selbst entscheiden, welchen Themen du dich zuwendest. Verbesserungen von kleinen formalen Änderungen bis hin zu größeren Verbesserungen hinsichtlich Quantität und Qualität sind denkbar. Dabei kannst du auf jegliche Quellen (Internet, Bücher, Weiterführungen aus Inferenzstatistik, Wissen von Bekannten, eigenes Vorwissen etc.) zurückgreifen. Achte bitte darauf, dass du nicht nur durch bloßes Einfügen das Wiki bearbeitest, sondern auch fließende Übergänge zwischen den einzelnen Aspekten herstellst, um so ein besseres Verständnis für die Zusammenhänge herzustellen. Die Auswertung der Wikis dient dabei lediglich meiner Bachelorarbeit und hat keinerlei sonstige Auswirkungen.

Bei der Zusammenarbeit ist außerdem von Bedeutung, dass du dich an folgende Vorgehensweise hältst: Mögliche Änderungen werden vor dem Bearbeiten zuerst den anderen über das jeweilige Diskussionsforum vorgeschlagen und diese dann in der Community diskutiert, bis eine Übereinstimmung bzw. ein Konsens erreicht ist. Erst dann erfolgt eine Bearbeitung der Artikel. Übereinstimmungen und Bearbeitungen müssen dabei allerdings nicht endgültig sein und können auch wieder verändert bzw. erneut im Diskussionsforum angesprochen werden. Die genaue Vorgehensweise, an die du dich halten sollst, ist oben (zum Vergrößern des Zyklus Grafik anklicken) und links abgebildet. Um die volle Versuchspersonenzahl zu erhalten, ist es wichtig, dass du dich 2,5 Stunden mit der Bearbeitung der Artikel auf dieser Plattform beschäftigst. Dies ist nachvollziehbar und wird vor Versuchspersonenvergabe kontrolliert. Dabei zählt nicht nur die Lese- bzw. Log-In-Zeit, du solltest versuchen, dich in jeden der oben und links abgebildeten Punkte einzubringen. 5 Log-Ins gelten dabei für dich als Richtwert, allerdings kann dies natürlich auch nach oben ausgeweitet werden. Um eine Zusammenarbeit zu gewährleisten, ist es wichtig, dass du nicht nur in einer der Wochen aktiv bist. Verteile die Log-Ins so, dass du ungefähr gleich viele Log-Ins in der ersten und zweiten Woche der Bearbeitung hast. Dies gibt dir die Möglichkeit auf neue Beiträge zu reagieren und ist bedeutsam für diese Studie. Schau dir außerdem die Artikel- und Diskussionsbeiträge anderer an und reagier auf diese entsprechend der vorgestellten Vorgehensweise. Ich würde dich darum bitten, nur auf dieser Plattform zu kommunizieren, um die Resultate nicht zu manipulieren.

The left navigation pane contains the following items:

- Navigation
- Vorgeschlagene Änderungen werden präsentiert
- Vorschlag diskutieren, temporären Konsens erreichen
- Artikel bearbeiten
- Konsens ist erreicht
- Vorgeschlagene Änderungen werden präsentiert
- Drucken/exportieren
- Druckversion
- Werkzeuge
- Links auf diese Seite
- Letzte Änderungen
- Medien-Manager
- Seitenindex
- Permanenter link
- Seite zitieren

Figure 4.7. Screenshot of one experimental group's wiki-based learning environment with collaboration script representations as used in Experiment 4.

4.4.3 Variables and Measurements

To measure the learning effect as one of the main dependent variables, two knowledge tests at different points in time were conducted. Prior to the wiki collaboration all participants received a 20-item knowledge test with five answer choices as the pre-test (e.g. "Which values are needed to calculate the confidence interval?"). Additionally, socio-demographic variables such as age and sex were also surveyed in the pre-test. After the collaborative writing phase participants answered a 20-item knowledge post-test that was structurally identical to the pre-test, also with five answer choices each (e.g. "Which conditions have to be satisfied to conduct a *t*-Test?"). Each wiki topic was addressed by two questions in both tests. Each question had one to three correct answer options. It was ensured that question difficulty, content areas, and the number of correct answer options per item were similar between the pre- and post-tests to create equivalent question sets. For both tests, the questions were generated using the original, unedited material in the wikis. The maximum number of points for both tests was set to the number of questions (20). Participants could get 0.2 points for both choosing correct answers and not choosing false answers. Although groups of students were surveyed, individuals have been chosen as unit of analysis, $ICC = .07$, $F(31, 32) = 1.16$, $p = .342$. Beyond that, wiki contribution data was collected by the learning environment, such as number of revisions per article, number of discussions per article, article word count and number of formal errors (e.g. spelling, grammar). For further analyses, several influencing variables were measured: need for cognitive closure (again measured with the German short scale 16-NCCS), intrinsic motivation and metacognitive strategies. To assess intrinsic motivation, the German short scale of intrinsic motivation (Wilde et al., 2009) was used, which is an adapted version of the Intrinsic Motivation Inventory (Deci & Ryan, 2003).

This scale has a four-dimensional factor structure and consists of the sub-scales Interest/Enjoyment (e.g., “The collaboration was enjoyable.”), Perceived Competence (e.g., “I think I was pretty good in the collaboration.”), Perceived Choice (e.g., “In the collaboration I could choose on my own how to handle it.”), and Pressure/Tension (e.g., “While collaborating I was tense.”) with three items in each subscale. Answers were given on a 5-point Likert scale from 1 = “strongly disagree” to 5 = “strongly agree”. The questionnaire to assess metacognitive strategies was taken from the inventory for the acquisition of learning strategies in studying (Wild & Schiefele, 1994). This is a 77-item questionnaire for the assessment of cognitive, metacognitive and resource-related learning strategies. As this study only focused on the metacognitive strategies the inventory was reduced to the eleven relevant items measuring this factor (e.g., “I ask myself questions about the material to make sure that I have understood everything.”) with answer options scaled from 1 (“very rare”) to 5 (“very often”).

4.5 Results of Experiment 4

4.5.1 Analytic Procedure

Due to the larger sample as in Experiment 3, more sophisticated statistical methods could be used to analyse the data. In addition to the direct two-group comparisons, moderation analyses were used to provide in-depth analyses of the conditional effects of intrinsic motivation, metacognitive strategies and the need for cognitive closure on contribution frequency during collaboration for both wiki script groups. If regions of significance were identified by the PROCESS macro applying the Johnson-Neyman technique, the conditional effect slope and its 95% upper and lower confidence bands were visualised (Hayes, 2013). Analogous to the analyses of

contribution frequency, moderation analyses were used to analyse the conditional effects of the discussed influencing variables on learning outcomes after collaboration for both wiki script groups. As before, if the PROCESS macro identified regions of significance with the Johnson-Neyman technique the appropriate plots were generated, otherwise simple slopes are presented for better visualisation purposes. As a final analytic step, potential links were explored between collaboration scripts and learning outcomes after collaboration, mediated by wiki contributions frequency.

4.5.2 Direct Group Comparisons

Learning outcomes. On a descriptive level, students in the DDR group performed higher in terms of absolute gains from the knowledge pre-test to post-test ($M = 1.04$, $SD = 1.43$) compared to the BRD group ($M = 0.91$, $SD = 1.54$), but the effect of the scripting was negligibly small, $t(63.94) = 0.35$, $p = .364$, $d = 0.08$, 90% CI [-0.31, 0.48], $BF_{01} = 1.14$. A paired samples t -Test suggested that students in both groups had learning gains from the first test ($M = 13.13$, $SD = 1.54$) to the second knowledge test ($M = 14.11$, $SD = 1.53$) with a medium to large effect for the time of testing, $t(68) = 5.52$, $p < .001$, $d = 0.66$, 90% CI [0.45, 0.89], $\log BF_{10} = 10.85$. Students mostly benefited from simply working with the wiki for a fixed period. Their learning gains did not immediately benefit from the collaboration script they were assigned to.

Edits and discussions. Regarding the wiki contributions, students in the DDR group participated more frequently in talk page discussions than students in the BRD group, $t(36.45) = 7.48$, $p < .001$, $d = 1.68$, 90% CI [1.17, 2.32], $\log BF_{10} = 13.56$. Students in the DDR group contributed fewer edits to the ten wiki articles than the BRD group, $t(63.09) = 2.11$, $p = .019$, $d = 0.51$, 90% CI [0.10, 0.91], $BF_{10} = 6.70$. These

contribution metrics results show that the collaboration scripts and their representations induced the intended collaborative knowledge construction behaviour for both groups. Wikipedia's BRD proposal promoted a rather high frequency of edits, whereas the alternative DDR script promoted more active engagement in discussions before meaningful changes to an article were made. Furthermore, it has been found that the BRD group added more words ($Mdn = 353.50$) to the original articles than the DDR group ($Mdn = 105.00$), $U = 87.00$, $p = .012$, $r = .58$, 90% CI [.24, .80], $BF_{10} = 7.79$. A positive correlation with a moderate to large effect between article length and the relative number of formal errors per article was also found, $\rho(18) = .48$, $p = .018$, 90% CI [.12, .72], $BF_{10} = 3.12$. These results indicate that the exhaustiveness of an article does not necessarily reflect its quality and that shorter articles are potentially more precise and accurate.

4.5.3 Conditional Effects on Contribution Frequency

Need for cognitive closure. There was a moderate to large conditional effect of the individual need for cognitive closure on the students' wiki contribution frequency, $F(3, 65) = 2.60$, $p = .060$, $R^2 = .10$, 95% CI [.00, .24], $BF_{10} = 15.79$ (Figure 4.8, left). Students with a high need for cognitive closure contributed more to the wiki when they were in the DDR script group. For students with a low need for cognitive closure the overall contribution frequency was not affected by the wiki script group they used. The DDR script encouraged those students, who would normally contribute less due to their disposition, to collaborate more.

Metacognitive strategies. There was a rather large conditional effect of the prevalence of metacognitive strategies on the contribution frequency, $F(3, 65) = 3.23$, $p = .028$, $R^2 = .13$, 95% CI [.00, .27], $BF_{10} = 39.44$ (Figure 4.8, middle). Students

who had less pronounced metacognitive strategies contributed substantially more to a wiki when they were assigned to the DDR script, whereas those with more pronounced metacognitive strategies did not differ in their contribution frequency between groups. The DDR script encourages contributions from students who need more instructions on what to do in a collaborative knowledge construction task.

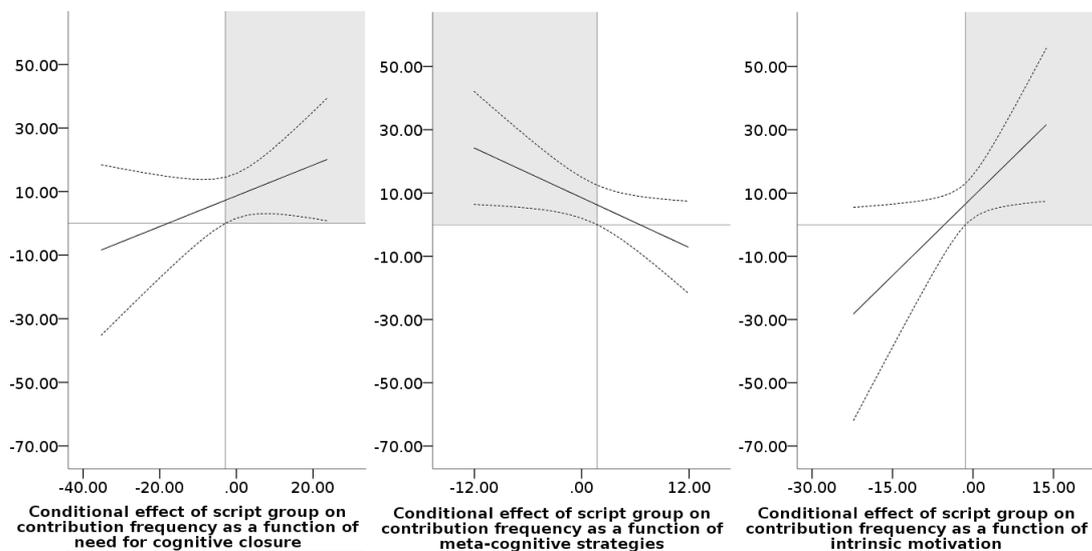


Figure 4.8. Conditional effects of script group on contributions as functions of influencing variables: need for cognitive closure (left), metacognitive strategies (middle), intrinsic motivation (right). Grey area = JN region of significance, dashed lines = 95% upper and lower confidence bands.

Intrinsic motivation. Finally, the analysis of intrinsic motivation suggests a large conditional effect on students' wiki contributions, $F(3, 65) = 3.29$, $p = .026$, $R^2 = .29$, 95% CI [.09, .45], $\log BF_{10} = 9.39$ (Figure 4.8, right). Students in the DDR script wiki contributed the most when their motivation was high. When motivation was low, the contribution frequency in the DDR script group showed a slightly reversed effect resulting in almost equally few contributions as in the BRD script wiki. Motivation seems to play a more important role for participation in collaborative knowledge construction activities if the deployed script is more restrictive.

4.5.4 Conditional Effects on Learning

For all three analyses of the influencing variables on learning gains from the pre- to the post-test, the Johnson-Neyman floodlight analysis did not identify any regions of significance. Thus, for the sake of completeness simple slopes as visualisations of the effects were provided. Analogous to the analyses on contributions conditional process analyses were performed with the three hypothesised moderating variables.

Need for cognitive closure. The first analysis suggested a small to moderate conditional effect of the need for cognitive closure on learning gains, $F(3, 65) = 0.87$, $p = .460$, $R^2 = .04$, 95% CI [.00, .13], $BF_{10} = 2.87$ (Figure 4.9, left). Students who were high on the closure scale had the largest gains in the BRD wiki and students who were low on the scale performed best in the DDR wiki. However, for both script groups and both extremes on the need for cognitive closure scale the average gains in test scores were relatively small.

Metacognitive strategies. Regarding metacognitive strategies as moderating variable, there was a small conditional effect on the learning gains, $F(3, 65) = 0.61$, $p = .610$, $R^2 = .03$, 95% CI [.00, .10], $BF_{10} = 2.19$ (Figure 4.9, middle). Having low levels of metacognitive strategies does not discriminate much between students of either wiki script, with a marginal advantage for those in the DDR group. Conversely, students who are high on the metacognitive strategy spectrum performed slightly better when collaborating in the BRD wiki.

Intrinsic motivation. Finally, regarding the conditional effects of intrinsic motivation on individual learning gains the analysis suggested there was a very small effect close to zero, $F(3, 65) = 0.10$, $p = .960$, $R^2 = .01$, 95% CI [.00, .03], $BF_{10} = 1.20$ (Figure 4.9, right). For students with high levels of intrinsic motivation it made

virtually no difference to which wiki script group they were assigned, resulting in almost identical learning gains. Students who had low intrinsic motivation for collaborating with other students in the wiki performed marginally better when they were working in the less strictly structured BRD wiki.

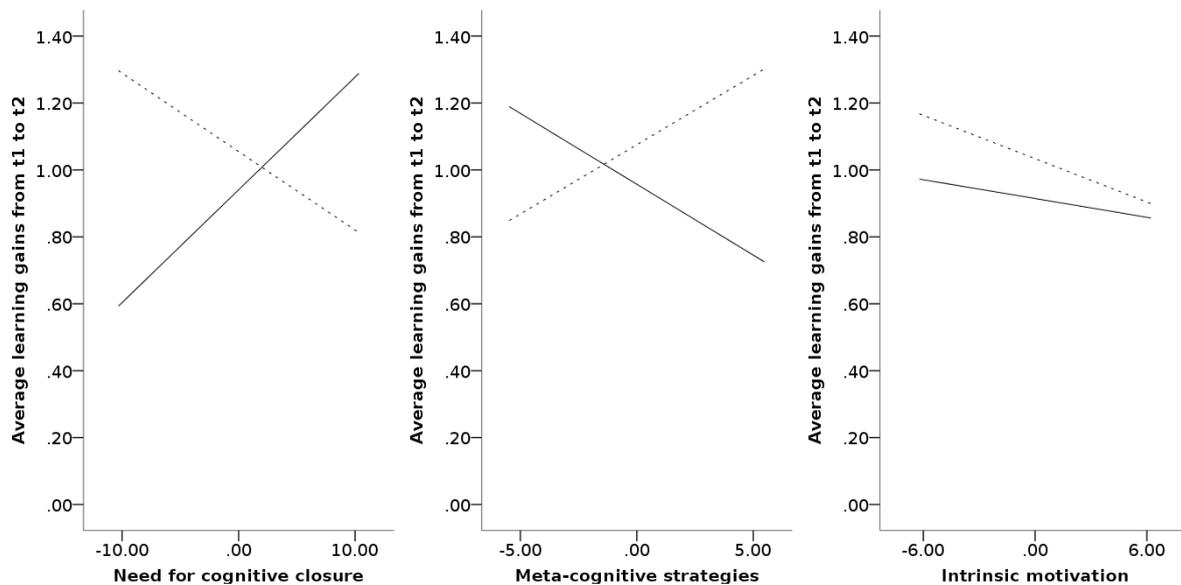


Figure 4.9. Simple slopes of script groups on learning gains conditional of influencing variables: need for cognitive closure (left), metacognitive strategies (middle), intrinsic motivation (right). Dashed lines = BRD script, solid lines = DDR script.

4.5.5 Mediation of Contributions Between Script and Learning

From a simple mediation analysis conducted using ordinary least squares path analysis, the provided collaboration scripts indirectly influenced the learning outcome at the post-test after collaboration (t2) through the students' wiki contribution frequency. As can be seen in Figure 4.10, students in the DDR wiki made on average more contributions to the wiki than students in the BRD group, $a = 8.55$, 95% CI [1.52, 15.57]. Thus, students who contributed more to the wiki were more likely to achieve a higher post-test score, $b = 0.02$, 95% CI [0.01, 0.04]. A bias-corrected

bootstrap confidence interval for the indirect effect based on 5,000 bootstrap samples was entirely above zero, $ab = .21$, 95% CI [.02, .52]. The total effect model for the unmediated direct effect highlights that the data do not sufficiently support a direct effect of the collaboration script on learning outcomes in the post-test, $F(1, 67) = 0.89$, $p = .349$, $R^2 = .01$, 95% CI [.00, .12], $BF_{10} = 1.29$.

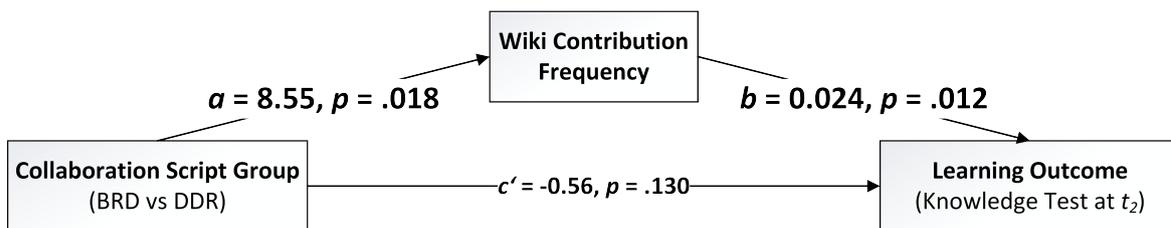


Figure 4.10. Effect of collaboration script on learning outcome mediated by the frequency of contributions to the wikis.

4.5.6 Further Results

Regarding the sentiments towards working with collaboration scripts and wikis, it was analysed if participants perceived the experimental wikis with their additional scripts as helpful. 67.57% of participants in the DDR group found that working with a wiki was generally useful compared to 71.88% in the BRD group, $\chi^2(1, N = 69) = 0.15$, $p = .698$, $BF_{01} = 3.45$. In the DDR group 83.78% of students perceived their provided script for collaboration with wikis as useful compared to 90.32% of students in the BRD group, $\chi^2(1, N = 68) = 0.63$, $p = .428$, $BF_{01} = 3.74$. In the DDR group 51.35% stated that they intend to use a wiki also in future works compared to 65.63% in the BRD group, $\chi^2(1, N = 69) = 1.44$, $p = .231$, $BF_{01} = 1.72$.

4.5.7 Summary of Main Findings

The central result of Experiment 4 is that students following different collaboration scripts in a real collaborative wiki environment show important differences in the resulting outcomes and underlying processes. Following the DDR script encouraged students to discuss substantial changes to articles before an edit was performed, whereas the BRD script group discussed virtually nothing and performed changes to the wiki whenever they felt the need. In most cases, students in the BRD wiki did not further proceed with the script's other steps of reverting and discussion. The script's aim is to promote many edits to an article; if no objections are communicated via revisions or discussions, it can be assumed that the community reached an implicit consensus about the performed edits. In contrast to that, in the DDR wiki it could be observed that many discussions between several students about the subject matter's subtopics took place. Accordingly, it can be assumed that the whole suggested script process was followed by most participants, since the whole process was observed of proposing an article edit in a discussion, followed by a deliberation process to reach consensus that finally resulted in a corresponding article edit. On the one hand, students produced longer articles in the BRD wiki, but on the other hand there were fewer errors in the DDR wiki articles. The greatest learning success at the post-test was most likely to be achieved when involvement in the wiki community was high, which was especially the case for students in the DDR script group. In extension to Experiment 3, further in-depth analyses were conducted of the influencing variables that thought to be related to learning and knowledge construction, namely need for cognitive closure, metacognitive strategies and intrinsic motivation. It has been shown that it is relevant to consider individual differences in these constructs when designing instructional aids for collaborative learning environments.

4.6 Discussion

Providing learners in technology-enhanced environments with explicit guidance on how to proceed during collaborative tasks can be highly relevant for more successful individual learning outcomes and collaborative knowledge construction artefacts. For the inception of new wiki articles or similar collaboratively created user-generated content, following an open script as proposed with Wikipedia's BRD can be useful. This script proposal encompasses the main stages of performing an article edit without any discussion (Be Bold), waiting to see if an edit will be reverted by another community member (Revert), and only take the edit as a topic to the discussion page if it gets reverted multiple times (Discuss). In contrast to that approach, when there is already a wiki knowledge base which can be built upon, it can be more useful to follow a script like DDR. That approach explicitly expects learners to coordinate and seek consensus before changes to the system are incorporated by promoting a discussion to present any planned edits (Discuss), finding a consensus with the community (Deliberate), and finally editing the article according to the consensus (Revise). Therefore, it was examined in detail how students work with either script in dyads and in larger groups. Differences in measurable learning outcomes were analysed and also the resulting artefacts that students produced in the experimental wiki environments. In the first experimental laboratory study with dyads, valuable first insights were gathered and evidence on how learners adopt different types of collaboration scripts. Moreover, it was examined how this affects their learning outcomes and knowledge construction in a wiki-like environment. Although the sample was quite small for reliable statistical inferences, some valuable qualitative information could be extracted from the learner's answers to the open-ended questions and some descriptive tendencies that provides some motivation to further explore differences of these two different script approaches.

For the most part, both scripts worked as intended. Students in the wiki environment with the DDR used the chat much more frequently for coordinating and discussing what content to include into the main article text. It could be seen that when students followed the suggested DDR script they were much more likely to incorporate their learning partner's additional information into more complex and nuanced answers. Although the number of well-integrated replies decreased over time, students who previously collaborated with the DDR script were still more likely to provide more nuanced answers than their BRD counterparts. Thus, collaboration with the DDR script caused students to add their own and the learning partner's arguments to reach more balanced articles. In the second study, it was built upon the findings of the first Experiment 3 and set up two real wikis for larger student groups as a supplement to a statistics lecture. Both collaboration script proposals worked well regarding their intended modes of operation. Students in the BRD wiki produced many more edits directly to the article than those who were in the DDR group. The main task in the BRD script is about editing an article and if no community member has any objections an implicit consensus is reached and no further script execution is required. In the DDR wiki students discussed most of their intended changes, whereas in the BRD group there were virtually no discussions during the experiment's observation period of two weeks. Students in the DDR wiki executed all stages of the entire script proposal by initiating discussions, participating in consensus finding and performing article edits as the last prompted step. The DDR script encouraged students to discuss on a level that is usually not present in educational wikis and more likely to be found in threaded online forums (Biasutti, 2017). It is highly appreciated that the students' adopted the DDR script steps as intended and thus resulting in increased knowledge exchanges about the subject matter in this group.

In contrast to the first experiment, the second study's set-up and subject matter made the occurrence of controversies less likely, making it more difficult to see learners' incorporation of complex and nuanced replies. Nevertheless, some important differences have been found. The resulting articles of the DDR wiki students were shorter, but they were more precise and contained fewer errors. In preliminary content analyses of a random sample of articles, it has been found that the articles produced in the DDR wiki were of higher overall quality (Heimbuch et al., 2016). Furthermore, the analyses of the influencing variables suggest that these findings are important to consider from an applied perspective. Other researchers of collaborative knowledge construction environments concluded that students can benefit from reading, connecting, and questioning ideas in an online environment (So, Seah, & Toh-Heng, 2010). Based on the study findings, it can be regarded as highly relevant to consider inter-individual differences when designing and incorporating guidance measures into computer-supported collaborative learning environments. Considerations of cognitive constructs such as the ones that have been presented can influence how students make use an online environment and what level of support might benefit them. In both experiments, effects of the individual need for cognitive closure were investigated. The results might seem ambiguous at first sight, partially because although the construct is regarded as a relatively stable disposition, it has also state component that can be influenced by time pressure (Webster & Kruglanski, 1994). The laboratory experiment had much stricter time limits of 45 minutes for the collaboration phase, whereas in the field experiment students had two weeks for collaboration with the wikis. As such, a high need for cognitive closure individual is likely to act differently under time pressure.

Regarding the additional two constructs that were assessed in the field experiment, metacognitive strategies and intrinsic motivation, it has been found that students in the more restrictive DDR wiki participated more than the average when they had few metacognitive strategies and were highly motivated. With high levels of metacognitive strategies, students do not necessarily need very explicit rules for effective collaboration, because they have already developed skills to easily adapt to specific requirements (Berthold et al., 2007; Kuhn, 2000). Finally, more intrinsic motivation can be required in more restrictive environments such as the DDR wiki, since the individual's autonomy of what to do is constrained by the suggested workflow. The idea of perceived autonomy is more in line with the concept of Wikipedia's BRD script which is less coercive and designed to incentivise high participation rates and many article edits (Nov, 2007; Zhang & Zhu, 2006). In conjunction, both studies suggest that when students collaboratively create user-generated content in environments like wikis, it is important to define the aims of collaboration and the whole setup. If the major goal is to gather as much information as possible in a short period of time and to fill up an empty knowledge base, open and less coercive guidance could be more suitable to achieve this. But if there is an already existing and established stock of knowledge to be improved or revised, it might be more useful to guide in a more restricted manner and to encourage discussions that are otherwise found in online forums rather than wikis. Moreover, as can be seen in the laboratory experiment, a more coercive script can induce more meaningful discussions and exchanges of arguments about opposing viewpoints. In future studies, it should be investigated how scalable collaboration scripts such as the DDR proposal are and if there is a critical mass of users that can efficiently make use of this.

5 The Interaction of the Need for Cognitive Closure with Implicit and Explicit Guidance in Wiki-based Learning

One purpose of wikis is the collaborative generation of content. During creation processes, controversies between authors emerge that they discuss on the article's talk page. Research suggests that controversies based on opposing points of view and contradictory evidence can be fruitful to trigger individual elaboration processes. However, previous research also showed that many wikis are not necessarily suited to identify relevant discussion contents and thus users need additional support as guidance. In an experimental study ($N = 181$) on wiki talk pages, two scaffolding measures were investigated in conjunction with the need for cognitive closure: (1) visual markers to highlight controversy status (implicit) and (2) a collaboration script that directs users towards discussions (explicit). Effects on wiki processes and learning outcomes were analysed. The results show that both guidance types can affect user behaviours and in interaction with the need for cognitive closure there are also effects on learning outcomes.

5.1 Method

5.1.1 Design and Participants

It has been decided to use an experimentally controlled laboratory setting with individual participants. This decision has been made to isolate potential effects of the experiment's guidance types in conjunction with the individual Need for Cognitive Closure from interfering effects caused by social interactions that naturally occur in wiki environments. Consequently, a between-subjects design was used to investigate the interplay between different guidance types and the Need for Cognitive Closure in wiki-based learning. The first independent factor was the type of provided talk page guidance (implicit vs. explicit). The second factor of interest was the individual Need for Cognitive Closure, which was factorised via median splits into two levels (low vs. high). For deeper inferential analyses, linear regression models were specified to make use of the full interval data spectrum. Before conducting the experiment, two a priori power analyses were performed corresponding to the main hypotheses. The first power analysis for between-group equivalence was performed with the R package *TOSTER* (Lakens, 2017a) for equivalence hypothesis testing with parameters $\alpha = .05$ $1-\beta = .90$, $d = [-0.5, 0.5]$. The bounds of $d = -0.5$ and 0.5 were chosen, because they translated into a raw score difference of approximately 1 point in the knowledge test, which was considered as the smallest effect size of interest (SESOI). The second power analysis for interactions of the grouping condition with the NCC was performed for linear regression designs in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) with parameters $\alpha = .05$, $1-\beta = .90$, $f = 0.25$. As a result, the analyses suggested an optimal $N = 174$ for the equivalence hypothesis tests and $N = 171$ for potential interaction effects in a hierarchical linear regression model.

After approximately five weeks, the recruitment of participants was terminated when no more subjects were willing to volunteer in the experiment. Finally, it was possible to sample $N = 181$ subjects with complete data sets. This sample size provides sensitivity for minimum equivalence bounds of $d = [-0.35, 0.35]$, which is more than sufficient for the study's planned bounds of $d = [-0.5, 0.5]$. The participants' age range was between 17 and 33 years ($M = 20.59$, $SD = 2.59$; $n_f = 136$ female, $n_m = 45$ male). Most of the participants were students recruited from the university's Applied Cognitive and Media Science degree program ($n = 168$; 92.82%). Participants were randomly assigned on their arrival at the laboratory to one of the two learning environments, resulting in an equal distribution to both an implicit guidance wiki ($n_{imp} = 91$) and an explicit guidance wiki ($n_{exp} = 90$). The participants' overall topic-specific interest was on a medium level ($M = 7.35$, $SD = 3.20$) and their self-assessed prior knowledge about the subject matter was relatively low ($M = 3.78$, $SD = 2.59$), on scales both ranging from "0 = low" to "15 = high". Differences between both wiki groups regarding topic-specific interest were very small, $U = 3917.50$, $p = .614$, $d = .04$, 95% CI $[-.13, .21]$, $BF_{01} = 5.28$. Regarding prior knowledge, a small but meaningful difference between group is suggested by the data, $U = 3359.50$, $p = .033$, $d = .18$, 95% CI $[-.02, .34]$, $BF_{10} = 2.16$. This difference in prior knowledge will be controlled for in respective analyses of the learning outcomes.

5.1.2 Procedure

The experiment was conducted in an individual setup with up to four participants at the same time, separated by divider panels. After participants were individually briefed with written instructions on the computer screen and had given consent to participate in the study, they were first asked a few basic socio-demographics as

well as interest in and prior knowledge of the study's subject matter (forms of energy). Participants completed all tasks of article editing and contributing to discussions individually in their own private wiki instances. This was followed by a short mandatory introduction to the self-developed wiki environment. Participants were asked to click through a mock-up environment with *lorem ipsum* texts to familiarize with the general wiki structure. In addition to the general orientation in a wiki, this tutorial phase also served to familiarize with the specific additions that were added to the experimental wikis (controversy highlighting vs collaboration script DDR) to ensure that participants have a common ground about their wiki environment's mechanics. Both groups had the same task of contributing to an initial Wikipedia-like base article about different forms of energy and participating in up to three of the corresponding discussions (cf. Figure 5.1).

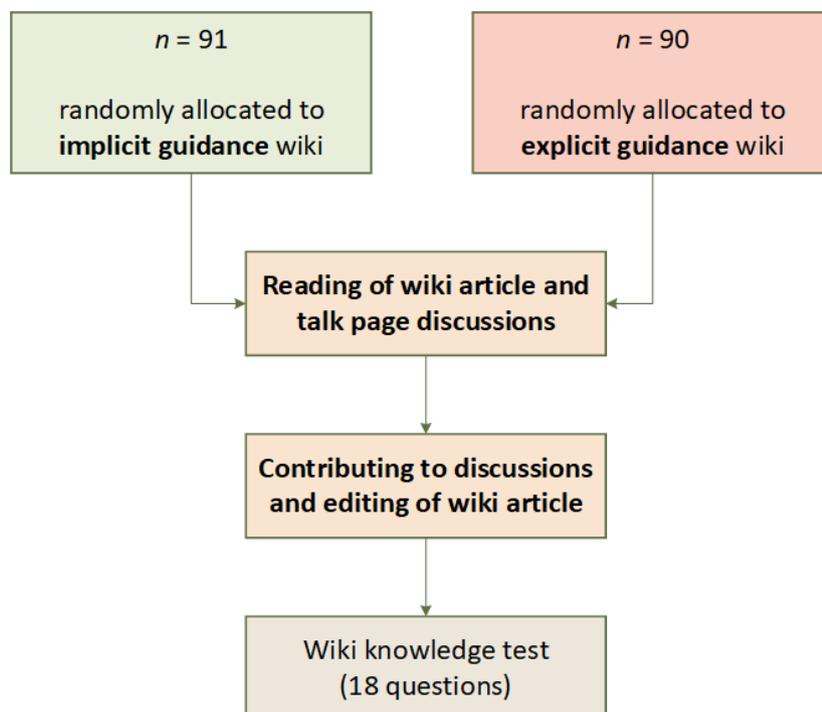


Figure 5.1. Workflow diagram visualising the overall study procedure of Experiment 5 with its central stages.

Participants received the information that the discussions contain sufficient arguments and evidence to enrich the original article, since no additional material regarding the subject matter was provided elsewhere. No further instructions were given on how to start their wiki task (e.g. reading the article or any discussion first) or what kind of reply they should make to a self-selected discussion. This was the experiment's main stage where participants had a loose total time limit of 21 minutes (three phases of 7 minutes for wiki contributions) for finishing all article edits and discussion replies. After the time for a contribution phase was up, the environment automatically prompted them to finish their contributions in the wiki and proceed further. Followed by the wiki contribution stage, the questionnaires to determine the individual levels of *Need for Cognitive Closure* (16-NCCS) and *epistemic curiosity* (ECS) were presented. After filling out these questionnaires participants had to answer a multiple-choice test about the study's contents. As an additional manipulation check, participants were asked to sum up briefly in open text fields why they have selected certain discussions to comment on and what led to the final decisions for the resulting article edits. Finally, to gain insights about how participants rate the additions made to the wikis they were asked to fill out the *User Experience Questionnaire* (UEQ).

5.1.3 Materials and Wiki Environments

Participants were confronted with different forms of energy, such as fossil fuels, nuclear power and renewable energy as the experiment's subject area. A base article on the topic was provided as an initial start page to provide a common ground for all participants in the experimental wikis. This article was derived from original sections of Wikipedia and adapted them for the study's purpose, resulting in an article with a total length of 630 words. From original talk page discussions on the

corresponding Wikipedia articles, a total number of twelve discussion threads were generated with the aim of reproducing a wiki-like environment (Figure 5.2). All threads in the wiki were made up of at least two discussants and included these on a wiki talk page to represent existing discussions that directly relate to the main article. Six of the integrated discussions comprised content-related controversies with opposing points of views about forms of energy and were resolved after the exchange of a few evidence-based arguments. The other six discussion threads were open and unresolved controversial discussions where discussants did not reach any kind of consensus.



Figure 5.2. The implicit guidance wiki with controversy highlights is represented in the floating upper left part of the figure. A DDR collaboration script representation for the explicit guidance wiki is shown on the right.

Since the DDR collaboration script was designed for discussion and deliberation of change proposals, it was necessary to simulate this step to a minimal degree for this individual study. It was decided to shorten the presented discussions for this group by one reply of a previous discussant and adapt it as boilerplate text for a simulated bogus discussant. If a study participant in the script group decided to reply to a self-selected discussion, a pre-selection of decisions was presented with three

options: (1) "I agree with discussant A", (2) "I agree with discussant B" and (3) "I agree with neither A or B / Both replies are equally valid to me." Depending on the user selection and the discussion status, one of three pre-defined bogus discussant replies was presented.

5.1.4 Variables and Measurements

The main independent factor of this study was the kind of additional wiki guidance, whether it was implicit with visual cues or explicit with a collaboration script suggesting a specific order of events. As another variable that also served for variance analytic purposes as a second independent factor was the individual Need for Cognitive Closure (NCC). According to the original literature on the scale 16-NCCS that has been deployed in this study, post-hoc median splits were used for classifying participants as low or high Need for Cognitive Closure (Schlink & Walther, 2007). Since median splits should only be used with caution due to false estimations, lower power and spurious statistical significance (Maxwell & Delaney, 1993; MacCallum, Zhang, Preacher, & Rucker, 2002), beyond the questionnaire's original evaluation protocol also the full continuous data spectrum of the Need for Cognitive Closure scale has been used within hierarchical linear regression models. On the other hand, median splits can also be useful to suggest clear recommendations to participants and to use these categories to build simple adaptive learning environments for different types of students. Thus, the following analytical procedure for process and outcome variables encompasses (1) analyses with dichotomised variables as well as (2) analyses with the entire metric data spectrum of the Need for Cognitive Closure scale.

Measuring learning success. To measure individual learning success about the study's subject matter, a post-experimental multiple-choice knowledge test was developed. Such tests are still widely used to quantify learning results in collaborative and individual settings (Kent et al., 2016). In total, the test comprised eighteen questions about different forms of energy, such as different types of renewable energy sources (e.g. "What is the efficiency of water?", "What are the negative effects of wind turbines?"). Six of these questions were designed to be answerable with only the information provided in the original base article. Therefore, they were practically solvable without having read any of the discussion threads. The remaining twelve questions were constructed in a way that exactly one question covered one of the controversial discussion topics. Every multiple-choice question had four answering options comprised of up to three distractors and at least one attractor. The test's overall sum of correct answer options was used as a general indication for individual learning success about the study's subject matter.

Measuring process variables. Both self-developed wiki environments were designed to record participants' selection behaviour by measuring individual clicks on the article and discussion tabs as well as on individual discussion thread titles. On the talk page, clicking on a title was necessary to select and expand a thread and thus unveil its contents. By design only one topic could be open for reading at a time and had to be collapsed by clicking again before proceeding to the next topic of interest. For further processing click counts that triggered only the expanding/opening events were recorded in the log. Furthermore, the environment recorded the times of events such as time spent on the article page, time spent on individual discussions and time spent to write a reply to a topic. Discussions' reading times were measured by calculating the differences between thread opening and closing times. If a topic

was opened and closed more than once, the environment also recorded cumulative reading times for each discussion thread. Correspondingly, for the article page the environment's logging system also recorded the overall time spent with the article.

Measuring further potential influences. Individual epistemic curiosity was measured with the Epistemic Curiosity Scale (Renner, 2006). This validated questionnaire with a total of 10 items measures the two dimensions diversive and specific epistemic curiosity. Each of the two subscales consists of five statements (e.g. 'When I learn something new, I like to learn even more about it.') and had to be rated on a 4-point scale ranging from "fully disagree" to "fully agree". In addition to that, participants were also asked to rate their further experience with wikis in a short self-developed questionnaire with six items on a 4-point scale ranging from "not correct at all" to "fully correct". Items covered questions regarding passive use of and active participation in wikis.

Measuring user experience. To measure the acceptance of the two guidance implementations which have been implemented in the experimental wikis, the *User Experience Questionnaire* (UEQ) was used (Laugwitz et al., 2008). This questionnaire consists of 26 items measuring the perceived usefulness of user interfaces and their implementations. The UEQ does not provide a total score of the user experience, instead the construct is made up of six dimensions that are considered individually. Overall, values smaller than $M = -0.8$ are considered negative evaluations, values between $-0.8 < M < 0.8$ as neutral and values greater than $M = 0.8$ as positive evaluations of a tool on the respective dimension.

5.1.5 Specific Hypotheses to Research Questions

RQ5.1: How do (1) implicit guidance with controversy awareness highlights and (2) explicit guidance with a DDR collaboration script affect processes and outcomes in wiki groups?

H5.1a: The individual selection and reply behaviour is expected to be equivalent for both wiki groups.

H5.1b: The individual contribution time is expected to be different between the guided wiki groups. Due to the nature of the DDR script, it is expected that participants in the explicit guidance group spent more time on contributing to discussions in comparison to participants in the implicit guidance group.

H5.1c: The individual learning success is expected to be equivalent for both wiki groups.

RQ5.2: How does the individual Need for Cognitive Closure influence processes and outcome variables related to learning in the case of (1) implicit guidance with controversy awareness highlights and (2) explicit guidance with a DDR collaboration script?

H5.2a: The interaction of the Need for Cognitive Closure and the implemented guidance in the wikis determines the individual selection and reply behaviour. It is expected that high NCC participants select and reply mostly to resolved controversies when their status is visualised as it is the case in the implicit guidance wiki. For low NCC participants, it is expected that they behave equivalently in both wikis.

H5.2b: The interaction of the Need for Cognitive Closure and the implemented guidance in the wikis determines the individual contribution time. It is expected that high NCC participants spent more time contributing when implicit guidance is present. For low NCC participants, it is expected that they spent more time with contributions when explicit guidance is present.

H5.2c: The interaction of the Need for Cognitive Closure and the form of structuring determines the individual learning success. It is expected that high NCC participants score higher in a knowledge test when implicit guidance was provided. For low NCC participants, it is expected that they achieve higher scores when explicit guidance was present.

5.2 Results

For the following analyses, equivalence hypothesis tests have been used with the TOST (two one-sided *t*-Tests) procedure. This testing procedure is very useful when a null or very small effect is expected and has in general greater statistical power for gathering evidence for or against the absence of effects. It applies two directional one-sided *t*-Tests against a priori specified lower and upper effect size or raw score bounds (cf. Chapter 1.9). Simultaneously, a classic NHST (Null Hypothesis Significance Testing) *t*-Test of differences from a null effect is applied. Within the TOST framework, the use of confidence intervals is more prevalent than *p*-values. Highly simplified, 95% CIs including zero correspond to non-significant *p*-values, whereas intervals excluding zero correspond to significant results (Lakens, 2017a).

5.2.1 Testing Equivalence between Guidance Groups

Topic selection. Participants in the implicit guidance wiki selected on average $M = 4.21$ ($SD = 2.75$) resolved topics in comparison to $M = 4.60$ ($SD = 2.86$) resolved topic selections in the explicit guidance wiki. The mean selection difference of 0.48 for resolved topics is equivalent in a 90% CI [-1.17, 0.21] and not significant in a 95% CI [-1.30, 0.34] within raw equivalence bounds of -1.40 and 1.40 resolved topic selections. Unresolved discussion topics have been selected on average $M = 4.84$ ($SD = 3.88$) times in the implicit guidance wiki and $M = 4.24$ ($SD = 3.21$) in the explicit guidance wiki. The mean difference of 0.59 topic selections between the guided wiki groups is equivalent in a 90% CI [-0.28, 1.47] and not significant in a 95% CI [-0.45, 1.64] within raw equivalence bounds of -1.78 and 1.78 unresolved topic selections.

Topic contribution times. In the implicit guidance wiki participants spent on average $M = 386.46$ ($SD = 246.55$) seconds on contributions to resolved controversial topics, whereas in the explicit guidance wiki $M = 481.11$ ($SD = 270.70$) seconds were spent on contributing to these controversies. This mean contribution time difference of 94.64 seconds between both groups is not equivalent in a 90% CI [-158.30, -30.98] and also statistically significant in a 95% CI [-170.62, -18.67] within raw equivalence bounds of -129.45 and 129.45 seconds. Regarding unresolved controversial topics, participants in the implicit guidance group spent $M = 436.21$ ($SD = 298.56$) seconds on contributions. In the explicit guidance wiki they spent $M = 341.37$ ($SD = 228.61$) seconds on contributing to unresolved controversies. The mean difference in contribution times for unresolved controversial topics of 94.84 seconds is not equivalent in a 90% CI [29.51, 160.17] and also statistically significant in a 95% CI [16.86, 172.82] within raw equivalence bounds of -132.95 and 132.95 seconds.

Topic reply frequency. Participants in the wiki with implicit guidance replied on average $M = 1.42$ ($SD = 0.92$) times to resolved controversies. In the explicit DDR script wiki, they replied $M = 1.80$ ($SD = 0.74$) times to these controversial discussion topics. The mean difference in reply frequency between groups of 0.38 replies is not equivalent in a 90% CI [-0.58, -0.18] and statistically significant in a 95% CI [-0.63, -0.14] within raw equivalence bounds of -0.42 and 0.42. Since the number of total discussion replies was fixed to $n = 3$, analyses of the reply frequency to unresolved controversies would be completely redundant. Subsumed, the results regarding the process variables provide evidence for the first hypothesis about the equivalence of implicit and explicit guidance wikis.

Knowledge test. In the knowledge test, participants in the implicit controversy highlight group scored on average $M = 15.84$ ($SD = 3.43$) in comparison to an average score of $M = 15.47$ ($SD = 3.60$) in the explicit scripting group. A mean test score difference between both guidance groups of 0.37 points is equivalent in a 90% CI [-0.50, 1.23] and not significant in a 95% CI [-0.66, 1.40] within raw equivalence bounds of -1.76 and 1.76 points (Figure 5.3). This result provides evidence for the equivalence in learning outcomes, meaning that no group outperforms the other.

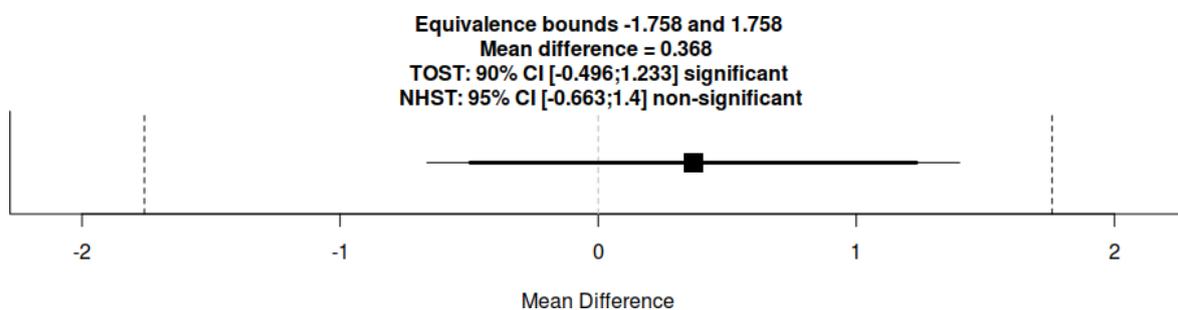


Figure 5.3. Equivalence hypothesis test using the two one-sided t -Tests (TOST) procedure between experimental groups on mean test scores.

5.2.2 Testing Equivalence between NCC Levels

Topic selection. Low NCC participants selected on average $M = 4.39$ ($SD = 2.59$) resolved topics in comparison to $M = 4.51$ ($SD = 3.04$) resolved topic selections of high NCC participants. The mean selection difference of 0.11 for resolved topics is equivalent in a 90% CI [-0.81, 0.58] and not significant in a 95% CI [-0.94, 0.72] within raw equivalence bounds of -1.41 and 1.41 resolved topic selections. Unresolved discussion topics have been selected on average $M = 4.35$ ($SD = 2.58$) times by low NCC participants and $M = 4.74$ ($SD = 4.36$) by high NCC participants. The mean difference of 0.39 topic selections between the guided wiki groups is equivalent in a 90% CI [-1.28, 0.49] and not significant in a 95% CI [-1.45, 0.66] within raw equivalence bounds of -1.79 and 1.79 unresolved topic selections.

Topic contribution times. Low NCC participants spent on average $M = 427.15$ ($SD = 245.71$) seconds on contributions to resolved controversial topics, whereas high NCC participants spent $M = 440.11$ ($SD = 279.94$) seconds on contributing to these controversies. This mean contribution time difference of 12.96 seconds between both groups is equivalent in a 90% CI [-77.78, 51.87] and not significant in a 95% CI [-90.33, 64.41] within raw equivalence bounds of -131.69 and 131.69 seconds. Regarding unresolved controversial topics, low NCC participants spent an average time of $M = 409.79$ ($SD = 298.79$) seconds on contributions. High NCC participants spent $M = 367.61$ ($SD = 235.40$) seconds on contributing to unresolved controversies. The mean difference in contribution times for unresolved controversial topics of 42.17 seconds is equivalent in a 90% CI [-23.83, 108.18] and not significant in a 95% CI [-36.61, 120.96] within raw equivalence bounds of -134.49 and 134.49 seconds.

Topic reply frequency. Low NCC participants replied $M = 1.65$ ($SD = 0.86$) times on average to resolved controversies. High NCC participants replied $M = 1.56$ ($SD = 0.85$) times to these controversial discussion topics. The mean difference in reply frequency between groups of 0.09 replies is equivalent in a 90% CI [-0.12, 0.03] and not significant in a 95% CI [-0.16, 0.34] within raw equivalence bounds of -0.43 and 0.43. Since the number of total discussion replies was fixed to $n = 3$, analyses of the reply frequency to unresolved controversies would be completely redundant. Subsumed for all process variables, the results provide further evidence for the hypothesis about the equivalence for low and high levels of need for cognitive closure.

Knowledge test. In the knowledge test, low NCC participants scored on average $M = 16.05$ ($SD = 3.29$) in comparison to an average score of $M = 15.24$ ($SD = 3.70$) of high NCC participants. A mean test score difference between both guidance groups of 0.82 points is equivalent in a 90% CI [-0.04, 1.68] and not significant in a 95% CI [-0.21, 1.85] within raw equivalence bounds of -1.75 and 1.75 points (Figure 5.4). This result provides further evidence for the expected equivalence in learning outcomes when only one factor is analysed in isolation, in this case the dichotomised (median split) need for cognitive closure.

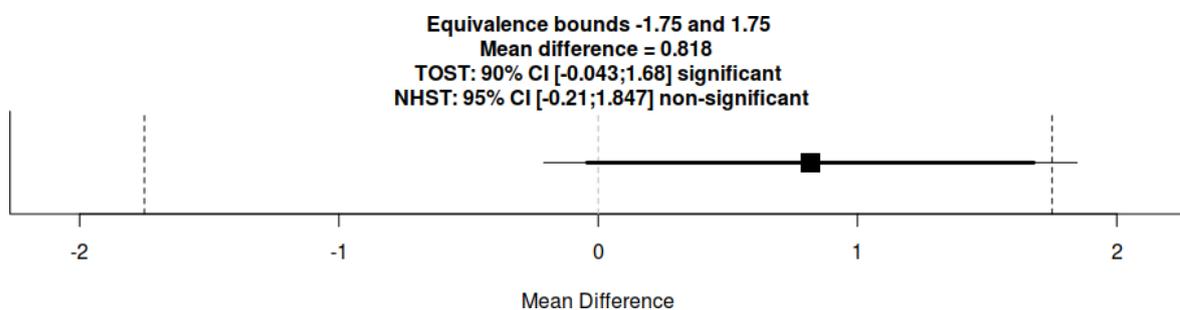


Figure 5.4. Equivalence hypothesis test using the two one-sided t -Tests (TOST) procedure between low and high NCC levels on mean test scores.

5.2.3 Testing Interactions of Guidance and NCC

In the following paragraphs on the analyses of interaction patterns, at first a 2 x 2 MANCOVA was used with the guidance type (implicit vs explicit) as group factor and the median split NCC (low vs high). Participants' self-assessed subjective prior knowledge was used as covariate in all the following analyses because substantial differences between the experimental groups have been identified. As subsequent analyses, hierarchical linear regressions were used to account for the full continuous data spectrum of the Need for Cognitive Closure scale.

Topic selection. In the MANCOVA model, the interaction effect of the experimental groups and the NCC level on topic selection behaviour was very small to virtually non-existent, $\lambda = .99$, $F(2, 175) = 0.83$, $p = .440$, $\eta_p^2 < .01$, 90% CI [.00, .04]. The follow-up regression analysis with the full continuous NCC data showed small effects for the interaction with the grouping variable on the selection of resolved topics with $b = 0.02$ ($SE = 0.04$), $t(176) = 0.46$, $p = .644$ within a small total effect model, $F(4, 176) = 0.72$, $p = .583$, $R^2 = .02$, 90% CI [.00, .04]. A similar pattern with small effects was found for the selection of unresolved topics with $b = -0.02$ ($SE = 0.05$), $t(176) = -0.47$, $p = .640$ in the total effect model, $F(4, 176) = 1.05$, $p = .385$, $R^2 = .02$, 90% CI [.00, .05].

Topic contribution times. For contribution times, the MANCOVA model showed an extremely small effect for the interaction of guidance type and the NCC, $\lambda = .99$, $F(2, 175) = 1.10$, $p = .335$, $\eta_p^2 < .01$, 90% CI [.00, .04]. Follow-up regressions with the NCC as continuous variable suggest some influence of the NCC interaction on the resolved topic contribution time with $b = 4.58$ ($SE = 3.61$), $t(176) = 1.27$, $p = .206$ within a medium-sized total effect model, $F(4, 176) = 3.03$, $p = .019$, $R^2 = .06$, 90% CI

[.01, .11]. A much weaker effect was found for the regression on unresolved topic contributions time of $b = 1.12$ ($SE = 3.70$), $t(176) = 0.30$, $p = .762$ that was also in a medium-sized total effect model, $F(4, 176) = 3.07$, $p = .018$, $R^2 = .07$, 90% CI [.01, .12]. The raw data suggests that prior knowledge is primarily responsible for differences in contribution times for both discussion topic categories.

Topic reply frequency. Due to the limitation of exactly $n = 3$ replies, instead of a MANCOVA on two reply variables a 2x2 ANCOVA has been used for analysing the interaction effect of guidance and NCC on resolved topic reply frequency. The effect with the dichotomized NCC was very small, $F(2, 175) = 1.91$, $p = .169$, $\eta_p^2 = .01$, 90% CI [.00, .06]. The regression model with the continuous NCC showed virtually the same pattern with a small effect with $b = 0.02$ ($SE = 0.01$), $t(176) = 1.32$, $p = .188$ within a rather larger total effect model, $F(4, 176) = 6.24$, $p < .001$, $R^2 = .12$, 90% CI [.04, .19]. As in the previous analysis on contribution times, the raw data suggests that prior knowledge is primarily responsible for differences in the replying behaviour. In summary, the analyses suggest relatively small to virtually no effects of the interaction between guidance and the NCC on all the process variables.

Knowledge test. Regarding the potential effects of the guidance type and the NCC level on the learning outcome, with a 2x2 ANCOVA a small to moderate effect was found, $F(2, 175) = 5.30$, $p = .023$, $\eta_p^2 = .03$, 90% CI [.01, .11]. Using the full NCC data spectrum with a hierarchical linear regression model, the effect of the interaction itself was weakened, $b = -0.03$ ($SE = 0.05$), $t(176) = -0.67$, $p = .502$ within a moderate total effect model, $F(4, 176) = 3.20$, $p = .014$, $R^2 = .07$, 90% CI [.01, .12]. The top and bottom panels in Figure 5.5 show the interaction diagrams according to the conducted variance and regression analyses. Both analyses provide evidence for the directions of the hypothesised effects.

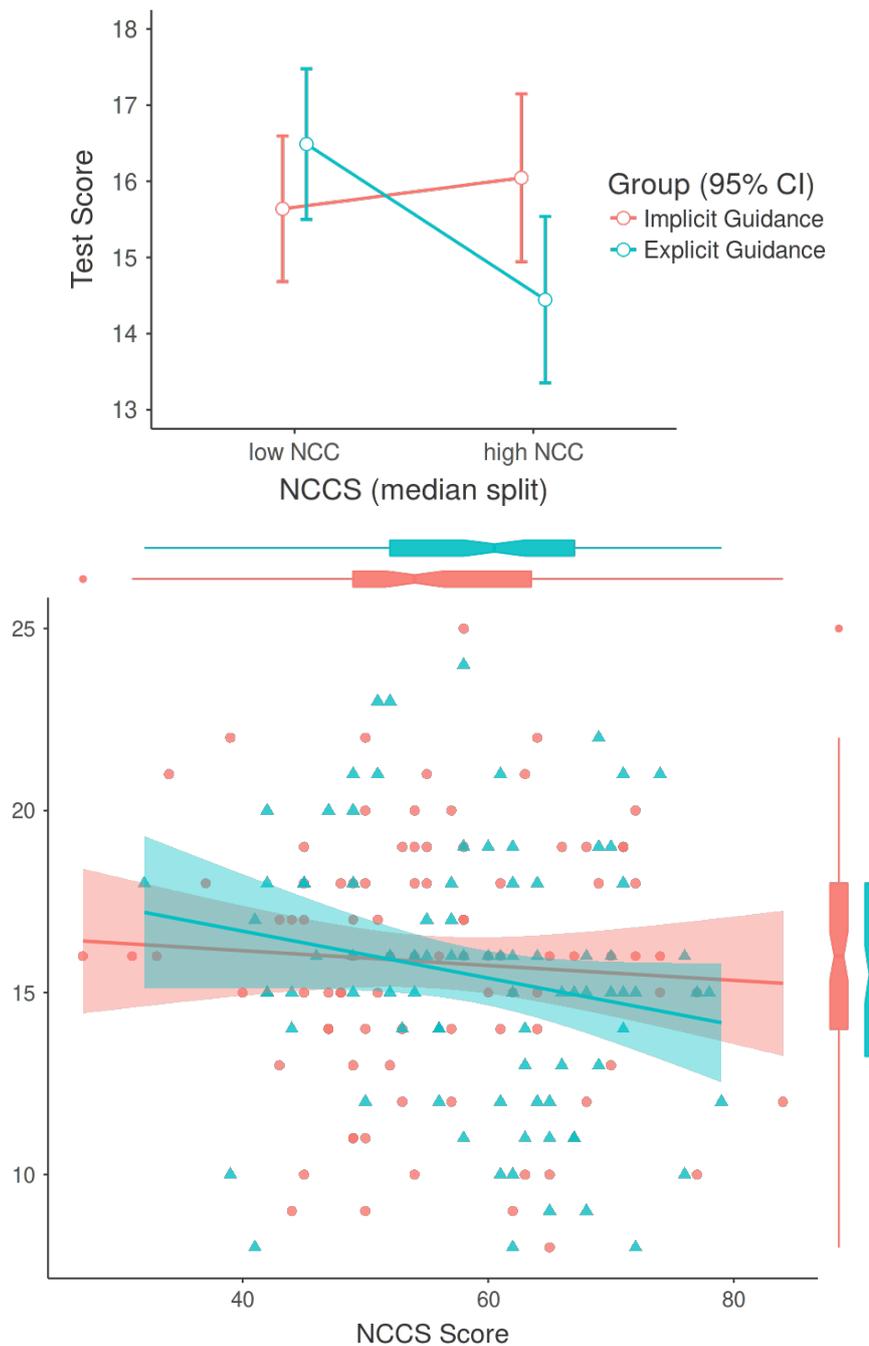


Figure 5.5. Interaction diagrams of the guidance type with Need for Cognitive Closure (NCC) on knowledge test scores. First, the interaction is visualised as in a 2x2 ANOVA with the dichotomised (median split) NCC (top panel). Second, the linear regression slopes are visualised with smoothed standard error areas for the continuous NCC spectrum (bottom panel).

5.2.4 Further Measurements

Epistemic curiosity. For the explored Epistemic Curiosity, there was a small positive direct effect on the knowledge test scores, $b = 0.09$ ($SE = 0.06$), $t(176) = 1.54$, $p = .124$. Participants scoring high on Epistemic Curiosity perform minimally better in the knowledge test. The overall effect model in the hierarchical linear regression was medium-sized, $F(4, 176) = 4.65$, $p = .004$, $R^2 = .07$, 90% CI [.01, .13]. A relatively large portion of the effect is due to differences in prior knowledge. The data of analysing the interaction of Epistemic Curiosity and guidance suggests a small positive effect on learning outcomes, $b = 0.15$ ($SE = 0.11$), $t(176) = 1.40$, $p = .164$. Participants scoring low on the used curiosity scale perform slightly better with controversy awareness highlights for implicit guidance. In contrast to that, participants on the higher end of the curiosity scale perform slightly better with the DDR script for explicit guidance. The overall effect model of the interaction hierarchical was medium-sized, $F(4, 176) = 3.99$, $p = .004$, $R^2 = .08$, 90% CI [.02, .14]. As before, some portion of the effect in the overall model is due to prior knowledge.

User experience. Overall, participants in the group with controversy awareness highlight for implicit guidance rated the status highlights in five out six UEQ dimensions higher than the collaboration script group for explicit guidance rated the DDR script. The largest difference on all six UEQ dimensions between both groups was on rating the *Efficiency* of their respective wiki environment, $t(179) = 6.44$, $p < .001$, $d = 0.96$. Students in the implicitly guided group rated their controversy status visualisations as rather positive in terms of efficiency whereas the explicitly guided experimental group gave even negative scores on the efficiency scale for the collaboration script. The only dimension were the explicitly guided group rated their wiki modification higher was the aspect of *Novelty*, $t(179) = -2.11$, $p = .037$, $d = 0.31$. The

detailed comparisons of all six UEQ dimensions between the experimental groups are shown in Figure 5.6.

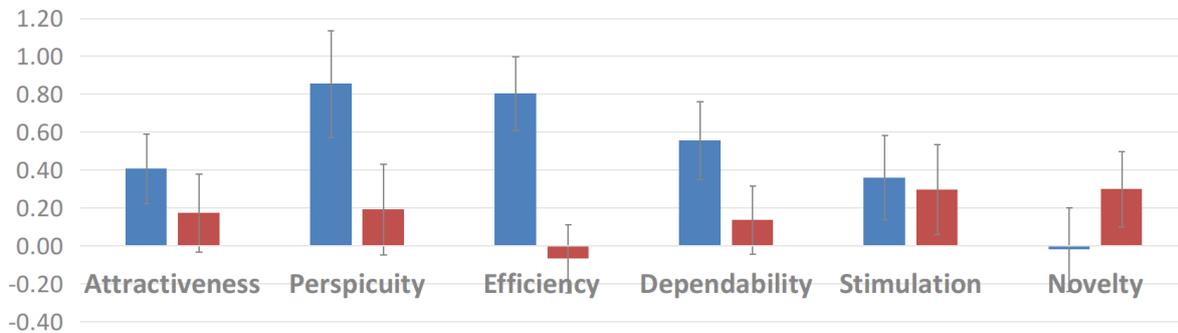


Figure 5.6. Mean ratings on the User Experience Questionnaire dimensions for controversy awareness highlights for implicit guidance implicit (blue) and collaboration script DDR for explicit guidance (red) implementations.

5.3 Discussion

Regarding the effects of modified wikis on activities, previous studies on implicit and explicit guidance have shown that such measures can influence behaviours of learners using wikis as learning environments. Implicit guidance with controversy awareness highlights can lead to a more focused selection of relevant content-related topics and explicit guidance with a discussion-centric collaboration script can lead to more meaningful a-priori discussion of proposed article changes (Heimbuch & Bodemer, 2015a, 2017). But in those previous research attempts individual cognitive differences of learning-related variables were mostly measured on the side. Therefore, the focus of this experimental study was laid on one of these constructs that had been previously identified as potentially relevant for learning scenarios with ambiguous information, namely the Need for Cognitive Closure (Webster & Kruglanski, 1994).

At first, the equivalence of implicit and explicit guidance in the overall sample was analysed, since there was no reason to believe that one measure in its own right should outperform the other. The results show that participants do not differ between the kind of guidance in the test scores. Regarding the process variables, they did not differ in the selection behaviour, but there were differences in the contribution times and replying behaviour. Participants with implicit guidance spent more time for their contributions when they were replying to unresolved controversies and contributed quicker and less frequent when dealing with resolved controversies compared to those with explicit guidance. Subsequently, participants were analysed regarding their individual Need for Cognitive Closure for equivalence with the construct dichotomised in low and high levels (Schlink & Walther, 2007; Schlink, 2009). There were no meaningful differences between participants in the two categorised levels of their Need for Cognitive Closure on any of the measured process variables or in the learning outcome. Finally, the interactions of the presented guidance measures and the Need for Cognitive Closure were analysed with (M)ANCOVAs and hierarchical linear regressions. For all process variables, identical patterns have been found with no differences in either statistical model. But when analysing the learning outcome, the interaction of guidance type and Need for Cognitive Closure was following the anticipated pattern. Participants who were high on the Need for Cognitive Closure scale scored higher in the knowledge test when working in a wiki enriched with implicit guidance rather than explicit guidance. The pattern reversed for low Need for Cognitive Closure participants, the scored highest in the discussion-centric collaboration script as explicit guidance wiki. Within a stronger statistical framework as with the used regression model the effects are weakened, but the general pattern remains about the same.

Persons with a high level of Need for Cognitive Closure have the desire to resolve ambiguity as quickly as possible and care less for the best possible solution and thus tend to rely on simpler heuristics to select and process information (Dijksterhuis et al., 1996; Webster & Kruglanski, 1994). Implicit guidance with the here user controversy awareness highlights provides them exactly this, a quick possibility to assess and to decide how to proceed further. In contrast to that, persons on the low end of the Need for Cognitive Closure prefer to elaborate in discussions and resolve ambiguity by finding better solutions than just the quickest (Dreu et al., 1999; Schlink, 2009). Thus, a more explicit guidance like the used discussion-centric script that tries to motivate the participation in discussion is better suited for these persons. Although the differences in raw test scores of approximately 1 to 1.5 are descriptively small, they can still be meaningful in the sense that even such a difference could be relevant for the next highest (or lowest) grade or can be a decision between passing and failing. Implicit and explicit guidance measures in socio-technical learning environments can have direct impacts on processes and even learning outcomes. But this study has shown that it is not necessarily the case. Other learning-related variables should also be considered to provide a more suitable and tailored learning experience to the individual prerequisites of learners. Some evidence was provided that individual differences in cognitive variables such as the Need for Cognitive Closure should be considered in designing and deploying learning environments. In addition, it can be noted that students in particular regard the controversy awareness highlights for implicit guidance on some dimensions as positive in terms of user experience. The visualisations are easily understood and their use is generally perceived as efficient. Only in the dimension of Novelty the explicit guidance through collaboration scripts was rated higher.

This may be due to the fact that most users are familiar with other visualisations of information, especially from the context of social media (e.g., Facebook, Twitter, Google+). Contrary to this, most of the participants have been confronted with collaboration scripts like DDR for the first time within the framework of this experimental study. As with most research conducted in higher education, this experiment was conducted with a very specific student sample of the Applied Cognitive and Media Sciences at a single university (University of Duisburg-Essen, Germany) that tends to be rather inclined to digital media. Therefore, it can be concluded that it would be very likely that the results are replicable within other populations of more than average technology-affine students. But it would be also very interesting to investigate the effects in populations that are less proficient with computers and socio-technical systems and to investigate if similar patterns would emerge.

Part III

General Discussion

6 General Discussion

Within the scope of this thesis the main question that was investigated was: “How can learners using wikis as learning environment can be supported to use the system beneficial for learning with the help of implicit and explicit structuring?” The research focus was not placed on the level of the wiki article, which emerges as a visible product of a collaborative writing and knowledge construction activity. Instead, the level of the wiki talk pages was examined in more detail, which serves as a forum in the background for coordination and discourse during the process of collaborative writing and co-evolution of knowledge. All the interventions that were analysed in Chapters 2 to 5 were designed and deployed considering the positive findings regarding implicit and explicit scaffolding developments in collaborative learning, both in face to face and in computer-mediated scenarios. One goal in the development, implementation and analysis of the implicit and explicit interventions was to make them as compatible as possible with current guidelines of existing wikis. The structuring aids as researched in this thesis should be easy to understand and use without the need for much in-depth training. Learners should be enabled to identify, apply and, at best, benefit from these measures.

6.1 Summary of Findings

The central results of all five experimental studies that were conducted during this PhD thesis project are summarised in the following. This includes a critical examination of the empirical work as well as the implications of the results for research and practice. Finally, further research areas relevant to the thesis topic of wiki-based learning as well as the significance of the results for future investigations will be discussed. The evidence of all studies' main findings is represented cumulatively within the model in Figure 6.1.

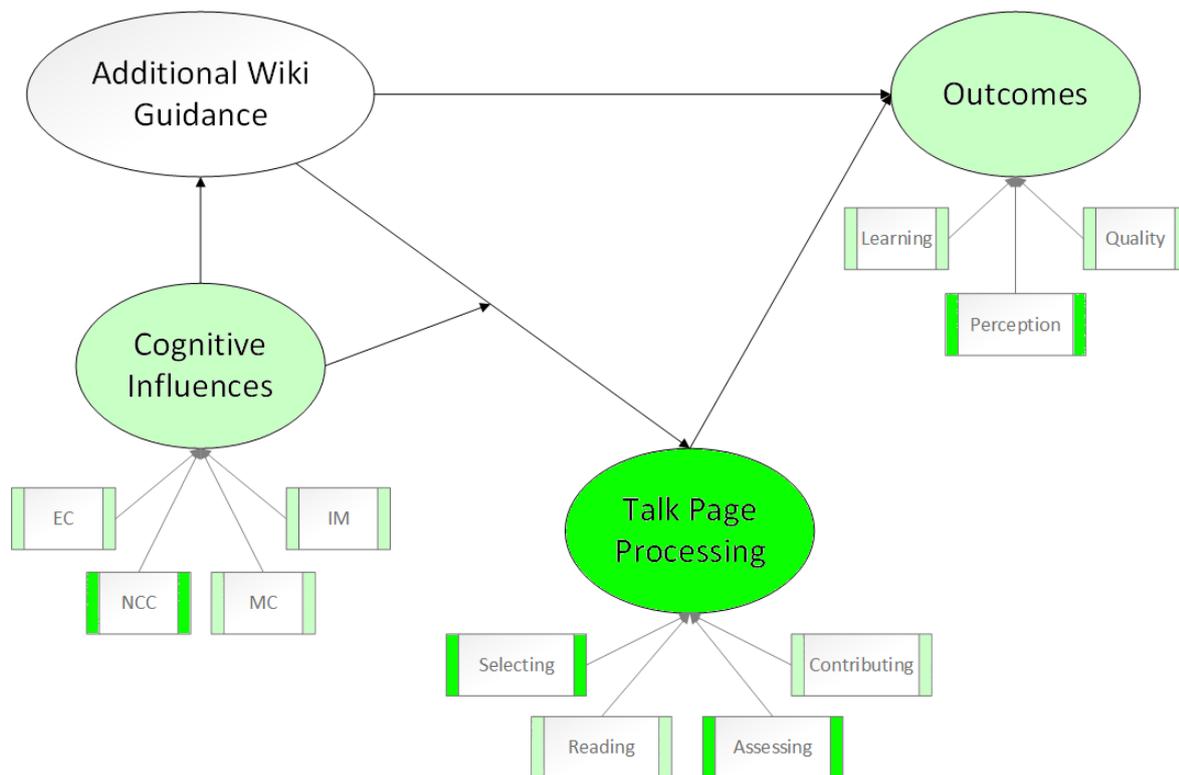


Figure 6.1. The direct and indirect influences of additional guidance measures on processes and outcomes. Light green represents moderate evidence for effects. Dark green represents stronger evidence for effects. The overall evidence for the constructs and variables is cumulative over all conducted studies.

Implicit guidance. It is particularly important to structure the interactions between learners to support a goal-oriented interaction and as a subsequent step also learning (Dillenbourg & Fischer, 2007). Within the scope of the experiments for this PhD project, study participants were directed by means of implicit structuring measures on the wiki discussion forums. In Experiments 1 and 2, and as well in Experiment 5, the effects of implicit structuring measures on different process variables were investigated. The analysed process variables were the topic selection behaviour (Experiments 1 and 5), discussion reading and contribution times (Experiments 1,2 and 5), topic reply behaviour (Experiments 1 and 5). Furthermore, different outcome variables were investigated in the studies, such as learning success in knowledge tests (Experiments 1 and 5), quality of article and discussion contributions (Experiment 1), subjective assessments of the perception of wiki talk pages and the expertise of authors (Experiment 2). In Experiment 1, study participants were directed with the help of controversy awareness highlights towards discussion topics that were relevant in terms of their relation to the article's content and provision of additional evidence. Such discussions can lead to socio-cognitive conflicts (Bodemer, 2011), which then lead to a more intensive examination of the topic and can thus be conducive to the learning process (Mugny & Doise, 1978). This experiment did not provide convincing evidence of directly measurable beneficial effects on learning outcomes or contribution quality. However, there were some positive tendencies with regard for the implicit structuring measures in wiki discussion forums. Specifically, when the process variables were examined more closely, there were some relevant effects of the investigated controversy awareness highlights. Participants who were provided with visual highlights were more likely to be attracted by highlighted discussion topics.

Because of that they were more likely to read the discussions more intensively and to reply to those controversial discussions as well. If this intended path was followed, higher test scores could be achieved, and higher quality wiki contributions were made. A different kind of implicit guidance was investigated with Experiment 2, with a study focus on expertise and credibility of wiki authors by means of social recommendation visualisations added to talk page discussions. There were no analyses of direct effects on learning outcomes, but of variables that can be indirectly related to learning processes as well. Credibility and trustworthiness of information sources can be crucial when people working on groups on the same product, as it is the case in collaborative knowledge construction (Lewis, 2003). As the experiment's results suggest, visualising author information on their expertise and social recommendations by the community can heavily influence the user perception. Although wiki users subjectively think that they are not influenced by this kind of information, objective data shows that they are. If expertise and community ratings are available, they are used as a guide to assess the credibility and validity of provided information and were helpful to identify content-related controversial discussions as such. In the two experiments 1 and 2, there were already some indications of possible influences by the individual Need for Cognitive Closure on knowledge construction processes and outcomes. Built upon these indications, in Experiment 5 interactions between guidance types and the Need for Cognitive Closure were examined more closely. An anticipated pattern on learning success could be found, where high Need for Cognitive Closure individuals who prefer to rely more on simple heuristics (Dreu et al., 1999; Kruglanski & Webster, 1996). The provided controversy awareness highlights for implicit guidance suit this need for a quick and simple measure more than an explicit collaboration script aiming at discussions and consensus building as it was intended with the DDR script.

Explicit guidance. Experiment 5 provided also evidence for the other side of the previously discussed interaction, regarding the interplay of explicit guidance with the individual Need for Cognitive Closure. When provided with a DDR script that proposes users to discuss wiki changes prior to their execution, users with a low Need for Cognitive Closure benefit from collaborating with this workflow resulting in more positive learning outcomes. These persons prefer to engage in discussions and enjoy dealing with ambiguity to generate more complex hypotheses and solutions to a problem (Dijksterhuis et al., 1996; Ford & Kruglanski, 1995). Although smaller in sample sizes, in both experimental studies on collaboration script, there were already some indications that individual differences in the Need for Cognitive Closure are correlated with differences in processing knowledge artefacts. This applies on the one hand to collaborative knowledge construction in dyads (Experiment 3), and on the other hand to collaborative knowledge construction in larger groups in a natural wiki environment that was used as a study course supplement (Experiment 4). The most remarkable result of the dyadic experiment was that participants who worked collaboratively with the DDR script were much more likely to integrate the evidence from their partner's additional learning material. The effect was most apparent when participants were surveyed immediately after the study, but even after a few weeks the effect remained in a follow-up survey. In Experiment 4, some additional individual differences were in the focus of analyses. There was some evidence that a more coercive collaboration script, such as the investigated DDR script, needs a relatively high intrinsic motivation in participants to be beneficial for the contribution quality and learning outcomes. Too many restrictions imposed by explicit guidance measures such as a DDR script can negatively influence the perceived sense of autonomy and thus can cause reluctance in participants (Deci & Ryan, 1985; Deci & Ryan, 2003).

Additionally, the individual availability of metacognitive strategies can be another relevant determinant for successful collaboration in a wiki as learning environment. When already a high degree of metacognitive strategies is accessible for a learner, then it does not really matter if the wiki collaboration is done after the rather open BRD script or the more explicit DDR script. These persons already have a set of elaborate strategies and are highly adaptable to different learning scenarios (Berthold et al., 2007; Kuhn, 2000). But when the degree of such strategies is low, which is often the case for novices, then it can be very beneficial for learners to be guided in more detailed steps through the entire process of knowledge construction with wikis.

6.2 Theoretical Implications

6.2.1 Implications for the co-evolution model

The presented experiments with a focus on learning and knowledge construction provide further evidence for processed as proposed in the co-evolution model for knowledge building (Cress & Kimmerle, 2008). The cumulative results of all knowledge tests support the assumption of internal assimilation to occur as part of individual quantifiable learning outcomes. Although the learning outcomes were not always observable as immediate effects leading back to the deployed interventions in the wiki environment, the evidence suggests that when learning-related processes are positively influenced that in the follow-up measurable differences in learning success can be observed (cf. Chapter 2). Although only one of the conducted studies explicitly investigated the process of internal accommodation, the evidence should not be neglected.

In Experiment 3 where dyads collaborated with contradictory additional learning materials, the results showed some remarkable differences in integrating and restructuring of one's own perspective on controversial aspects of the subject matter (cf. Chapter 4). When dyads collaborated with the rather open and more advising than prescriptive BRD script, it was more likely that no internal accommodation has taken place. Improvements in qualitative individual learning were mostly observable when dyads collaborated with a DDR script, proposing to discuss more intensively a priori about planned changes. This positive effect remained relatively stable when participants were surveyed repeatedly after a few weeks. Both processes of external assimilation and external accommodation could be observed in all presented studies where participants were asked to add new information or had the task to alter existing articles with evidence found in talk page discussions (cf. Chapters 2,4 and 5). These outcomes of the processes were most remarkably observable in the field Experiment 4 where larger groups of students collaborated in two real wiki environments and were either assigned to the BRD or the DDR script (cf. Chapter 4). When students worked with Wikipedia's proposal of BRD, it was most likely to observe quantitative knowledge building as external assimilation. They mostly added new information to the wiki articles without discussions or major restructuring of articles. Students who worked in the DDR script wiki, on the other side, were much more likely to also engage in qualitative knowledge building as external accommodation. Students were much more involved in discussing further wiki contributions and consequently these discussions led to shorter but more precise and less error-prone articles than in the BRD script wiki. The presented results support assumptions of the processes and effects of the co-evolution of knowledge building. Yet, they have not been fully systematically studied against the background of

the effects of implicit and explicit guidance. It remains yet unknown which kind of guidance measure can support which processes to what extent.

6.2.2 Implications for implicit guidance

Both kinds of implicit guidance measures in this project were focussing on distinct stages in the help-seeking process, namely stage 1 “Determine that a problem exists” and stage 5 “Decide whom to ask” (Makara & Karabenick, 2013). Stage 1 was mainly covered by the research conducted in Experiment 1, where participants were provided with additional controversy awareness highlights to signal the potential task relevance (cf. Chapter 2). These kinds of measure acting on the information as source level can be helpful for users who already have the desire to learn more about a topic and feel lost in the mass of information that can be found on some discussion pages. If the information on the controversy of a certain topic is not easily accessible, except for reading every available discussion, controversy awareness highlights can be a valuable information to reveal to an interested user that there is probably more to a subject matter than the wiki article initially suggested. Stage 5 was experimentally investigated in Experiment 2 when the effects of additional author information about their expertise and community-standing were in the focus of analyses (cf. Chapter 3). In this study’s scenario the initial situation for a wiki user was that one was already interested or even involved in a controversial discussion, but felt the urge to easily assess which arguments a discussant provided as an information source are trustworthy. When additional information about the author’s expertise and social recommendations by the community were present in talk page discussions, this lead users to use this information as primary guide for assessing the source credibility and consequently provided compelling evidence for the potentials of using such information as main assessment cues. Further stages

in the help-seeking framework were not addressed in any of the presented studies, neither implicitly nor explicitly. This would require more systematic investigations of the modes of operation of scaffolding measures within the help-seeking process with educational technology.

6.2.3 Implications for explicit guidance

The investigated scripts BRD and DDR in Experiments 3 and 4 add to the body of evidence regarding the potential effectiveness of collaboration scripts, especially when considering differences in script coercion (Papadopoulos et al., 2013; Wichmann & Rummel, 2013). A more coercive script such as the here investigated DDR script with a focus on a priori discussions might be regarded as less desirable by students, but the evidence shows that it can have beneficial effects for them. The BRD script leads to more article edits as expected and requires less coordination to perform any planned changes. Additionally, it also lead to longer articles in comparison to those edited within the DDR scripted wiki. The DDR script successfully promoted discussions between students before wiki changes were performed. This lead to more condensed articles of higher quality. And consequently, students who collaborated with the DDR script performed slightly better in a knowledge test (cf. Chapter 4). Both collaboration script approaches can be valid choices for different use cases. The BRD script can be an appropriate choice when the task is to create an entirely new knowledge base. Such an open workflow can promote a high number of knowledge externalisation, which would be desirable when a wiki is empty at first. As soon as a critical mass of information is reached in a wiki, it could be more beneficial for users to switch to a more coercive script like DDR, to further improve the knowledge base that is already there. But it is yet not clear when a critical mass is reached and thus would require more systematic research.

6.2.4 Further implications

When considering these individual differences in collaborative knowledge construction settings with educational technology like wikis, a proposal for a unified model of help-seeking (Makara & Karabenick, 2013) and knowledge building with socio-technical systems (Cress & Kimmerle, 2008) (cf. Figure 6.2).

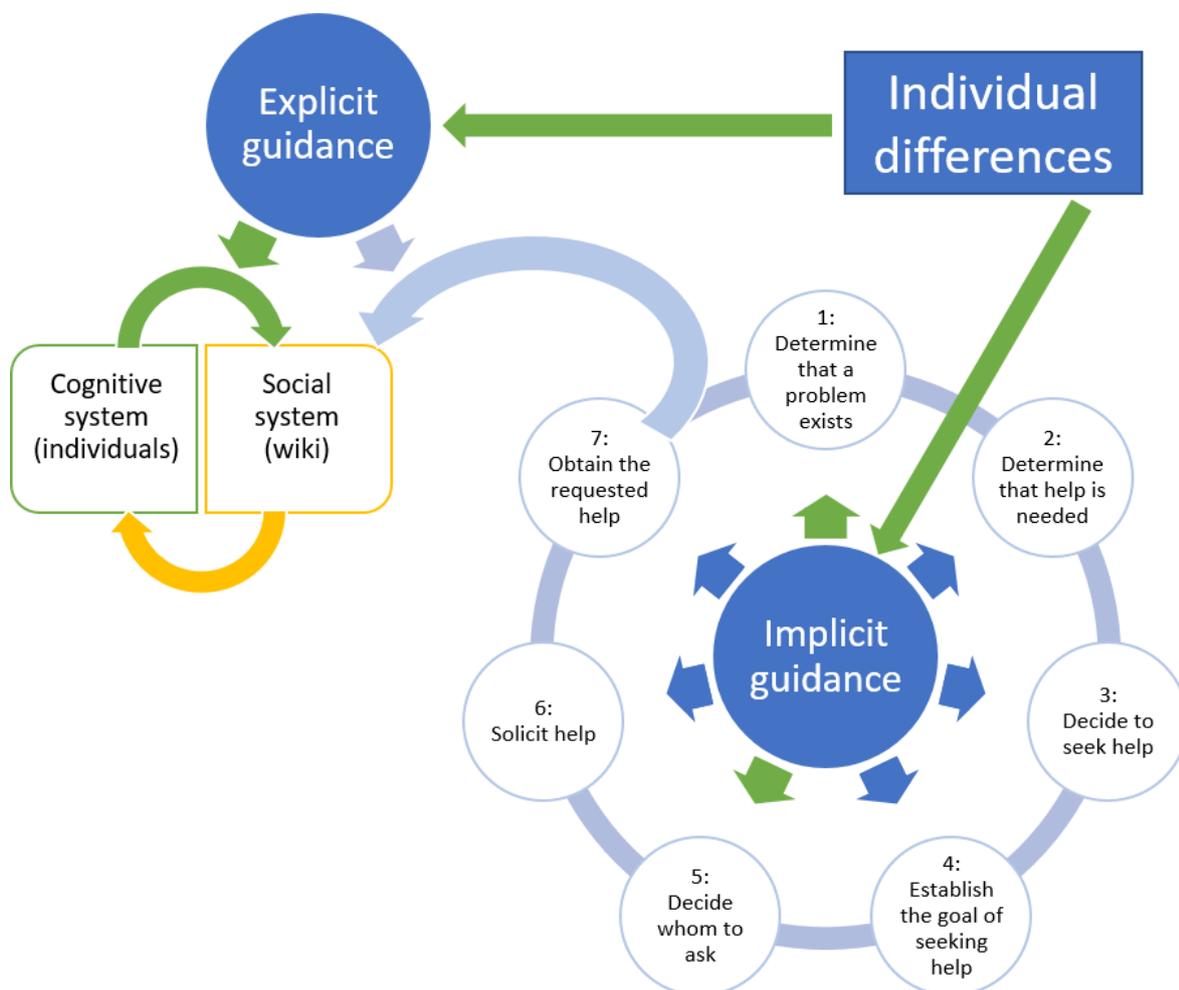


Figure 6.2. Proposal of a unified model for collaborative knowledge construction with wikis. Implicit guidance can be used to influence the help-seeking process at different stages. Going through this process can lead to contributions in the wiki that are determined by interchanging processes of internalisation and externalisation which can be influenced by explicit guidance. The effectiveness and efficiency of both types of guidance can be influenced by individual differences.

All experimental studies, although to different extents, provided evidence for the relevance of considering individual differences regarding influencing variables. Specifically, the relevance of the Need for Cognitive Closure in collaborative knowledge construction was emphasised because of the high likelihood of controversies and socio-cognitive conflicts to occur out of ambiguous information and during discussions of opposing evidence about a subject matter. Thus, it seemed inevitable to consider this individual construct that deals with preference and aversion of ambiguity, as states and traits, and is closely related to other constructs like Epistemic Curiosity that can also be seen as a motivational construct relevant for learning-related processes like information-seeking and interest. These individual variables influence the effects of additional guidance measures that are aimed at either supporting collaborative knowledge construction or specific stages of individual help-seeking for learning with wikis (cf. Chapter 5). There is yet not enough or no data available on all possible interaction effects of all the assessed individual variables and guidance types at all process stages that can be tied to collaborative knowledge construction and learning with wikis, and thus requires more systematic research.

6.3 Practical Implications

For the application of wikis in educational institutions, at work or in private contexts, it can be said that wiki platforms can be a suitable medium for supporting collaborative learning and knowledge construction. Although the studies conducted as part of this thesis did not show immediate effects on outcome variables when thinking of individual learning.

Some tendencies could be shown at a descriptive level, which speak in favour of the use of wikis, but also of considering additional structuring measures as scaffolds. In educational settings such as in exam situations or graded assignments, for example, even minor differences of just one or two points can be decisive for the passing or failing of an individual. It should be noted that both guidance types were analysed in settings where an already existing knowledge base had to be improved and refined. The appropriate structuring method must be chosen for the respective learning scenario to achieve positive results and thus the desired effect, as both structuring modes can be advantageous for the individual depending on the situation. For instance, rather simple and perspicuous structuring aids like the controversy awareness highlights for implicit guidance can be more efficient for a significant number of people, since a general overview of the relevant topics is possible in the given time and more knowledge could be acquired in a short time. This can be rather useful for projects where a fixed deadline is set. In terms of quality, on the other hand, the more complex structuring like a DDR script for explicit guidance can be more effective for the learning-related outcomes, since the learners were enabled and called upon to deal more intensively with the relevant content and thus gain more detailed knowledge about opposing points of view and contradictory evidence, which again had a positive effect on the quality of wiki contributions. Overall, it can be concluded that additional structuring of wiki discussions as learning material can have a positive effect on the learning success and the quality of contributions. With this knowledge, the present studies expand the field of empirical investigations in an educational context. At the same time, however, it points to a continuing need for research and provides important impulses for subsequent investigations.

In order to arrive at well-founded statements about the effect of implicit structuring measures in Wiki-based learning environments, studies with larger numbers of cases have to be carried out. A further approach could be to focus subsequent research on the influence of personality variables, which may have an impact on the learning effect and the quality of contributions. Such approaches could help to achieve more comprehensive results in the field of educational research and to derive beneficial strategies for instructional practices.

6.4 Outlook and Future Directions

Based on the results of the studies conducted within the course of this PhD project and with a view to future research, it can be stated that although it is important to choose the correct structuring method for the respective learning situation, this does not necessarily guarantee that the desired effects will occur. An important aspect is the specialisation of further research on individual variables, which, as can be seen from the discussed results, can have a decisive influence on the effects depending on the kind of added structuring. Although some results of the conditional analyses were not always conclusive, some descriptive tendencies were also evident here. More well-powered studies could help to draw more reliable conclusions about conditional process effects. Future research should therefore pay more attention to the interaction of these two types of structuring. For practical use, it is important to add that instructors should ensure that the collaboration meets the individual requirements of the learner, including motivational processes and individual differences. It should also be noted that most of the subjects in all here presented studies were not regular or usual wiki users.

The majority of participants was recruited as convenience samples out of a rather special student population, namely students of the Applied Cognitive and Media Science from the University of Duisburg-Essen in Germany. This degree programme is in that regard special because it combines studies of Psychology as well as Computer Science to a similar degree. Therefore, further research would be of interest to investigate whether actual Wiki(-pedia) users and as well more diverse student populations from different higher education programmes would benefit from the here investigated implicit and explicit guidance measures to the same extent.

6.5 Final Conclusion

The prevalent use of socio-technical systems like wikis for learning requires more advanced understanding of how people make use of these platforms to collaboratively construct knowledge and how they can be further supported to use them even more effectively and efficiently. For this purpose, the research conducted as part of this PhD project investigated different scaffolding measures of implicit and explicit guidance in light of the theories of co-evolution of knowledge and framework of help-seeking for educational technology. A crucial part of this research was whether wiki users understand, willing to adopt and make use of the deployed measures and thus can produce better knowledge artefacts and are enabled to learn more about a given subject matter. The series of five presented empirical studies provides a lot of compelling and as well some yet rather inconclusive evidence. While the effects on directly measurable process variables are relatively strong and clear, the effects onto learning outcomes and constructed knowledge artefacts are less clear. The results show almost no direct effects of additional guidance measures on outcomes, but some indications of indirect effects when considering the line of process variables.

When the guidance was accepted by the user and the intended steps were followed, there was some potential of these interventions to be beneficial for the individual learner as well as for the quality of the collaboratively constructed knowledge artefacts in a wiki. The final discussed experiment provides evidence for the interplay between learning-related cognitive variables and the kind of guidance measure that is implemented. Depending on one's personal level of Need for Cognitive Closure it can make a difference for the learning success if one has to use a wiki system with more or less restrictions regarding the knowledge construction process and more or less intuitive tacit guidance. The findings with regard to the other assessed influencing variables were less conclusive and raised more questions about their systematic effects on learning and knowledge construction outcomes. Nevertheless, the research presented here fits into a series of wiki-related educational psychology research, which previously focused mainly on effects at the article level of wikis, while the core of this research was mainly focused on shedding some light on the background level of wiki article creation, namely the talk pages.

References

- Agasta Adline, A. & Mahalakshmi, G. (2012). A recommender system for wiki pages: Usage based rating approach. In *2012 International Conference on Recent Trends In Information Technology (ICRTIT)* (pp. 553–558). <https://10.1109/ICRTIT.2012.6206831>
- Alamargot, D. & Chanquoy, L. (2001). *Through the models of writing* (G. Rijlaarsdam, Ed.). Studies in Writing. Dordrecht: Springer Netherlands.
- Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help Seeking and Help Design in Interactive Learning Environments. *Review of Educational Research*, 73(3), 277–320. <https://10.3102/00346543073003277>
- Bär, D., Erbs, N., Zesch, T., & Gurevych, I. (2011) (pp. 74–79). Stroudsburg, PA, USA: Association for Computational Linguistics.
- Bartelsen, J. & Brauer, J. (2010). *Kooperatives Lernen mit einem Wiki*. Arbeitspapiere der Nordakademie.
- Baumeister, A. E. E., Engelmann, T., & Hesse, F. W. (2016). One task, divergent solutions: High- versus low-status sources and social comparison guide adaptation in a computer-supported socio-cognitive conflict task. *Educational Technology Research and Development*, 1–17. <https://10.1007/s11423-016-9466-1>
- Beers, P. J., Boshuizen, H. P. A. (, Kirschner, P. A., & Gijsselaers, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior*. Learning in Innovative Learning Environments, 21(4), 623–643. <https://10.1016/j.chb.2004.10.036>

-
- Bell, N., Grossen, M., & Perret-Clermont, A.-N. (1985). Sociocognitive conflict and intellectual growth. *New Directions for Child and Adolescent Development*, 1985(29), 41–54. <https://10.1002/cd.23219852905>
- Bell, P. (2004). The educational opportunities of contemporary controversies in science. In *Internet environments for science education*. Mahwah, NJ: Lawrence Erlbaum.
- Berlo, D. K., Lemert, J. B., & Mertz, R. J. (1969). Dimensions for Evaluating the Acceptability of Message Sources. *Public Opinion Quarterly*, 33(4), 563–576. <https://10.1086/267745>
- Berlyne, D. E. (1954a). A theory of human curiosity. *British Journal of Psychology. General Section*, 45(3), 180–191. <https://10.1111/j.2044-8295.1954.tb01243.x>
- Berlyne, D. E. (1954b). An Experimental Study of Human Curiosity. *British Journal of Psychology. General Section*, 45(4), 256–265. <https://10.1111/j.2044-8295.1954.tb01253.x>
- Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17(5), 564–577. <https://10.1016/j.learninstruc.2007.09.007>
- Biasutti, M. (2017). A comparative analysis of forums and wikis as tools for online collaborative learning. *Computers & Education*, 111, 158–171. <https://10.1016/j.compedu.2017.04.006>
- Biasutti, M. & EL-Deghaidy, H. (2012). Using Wiki in teacher education: Impact on knowledge management processes and student satisfaction. *Computers & Education*, 59(3), 861–872. <https://10.1016/j.compedu.2012.04.009>

- Bobadilla, J., Serradilla, F., & Hernando, A. (2009). Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems. Artificial Intelligence (AI) in Blended Learning(AI) in Blended Learning*, 22(4), 261–265. <https://10.1016/j.knosys.2009.01.008>
- Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, 27(3), 1079–1086. <https://10.1016/j.chb.2010.05.016>
- Bodemer, D. & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, 27(3), 1043–1045. <https://10.1016/j.chb.2010.07.014>
- Bodemer, D., Gaiser, B., & Hesse, F. W. (2011). Kooperatives netzbasiertes Lernen. In P. Klimsa & L. Issing (Eds.), *Online-Lernen: Planung, Realisation, Anwendung und Evaluation von Lehr- und Lernprozessen online* (2nd ed., pp. 151–158). München, Deutschland: Oldenbourg.
- Bradley, M. M. & Lang, P. J. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In *Handbook of Emotion Elicitation and Assessment* (pp. 29–46). Oxford, UK: Oxford University Press.
- Bromme, R., Hesse, F. W., & Spada, H. (Eds.). (2005). *Barriers and biases in computer-mediated knowledge communication - and how they may be overcome*. Computer-Supported Collaborative Learning Series. New York, NY: Springer.
- Brown, M. B. & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346), 364–367.
- Buder, J. & Bodemer, D. (2008). Supporting controversial CSCL discussions with augmented group awareness tools. *International Journal of Computer-Supported Collaborative Learning*, 3(2), 123–139. <https://10.1007/s11412-008-9037-5>
- Buder, J., Schwind, C., Rudat, A., & Bodemer, D. (2015). Selective reading of large online forum discussions: The impact of rating visualizations on navigation

-
- and learning. *Computers in Human Behavior*, 44, 191–201. <https://doi.org/10.1016/j.chb.2014.11.043>
- Calvo, M. G. & Lang, P. J. (2004). Gaze Patterns When Looking at Emotional Pictures: Motivationally Biased Attention. *Motivation and Emotion*, 28(3), 221–243. <https://doi.org/10.1023/B:MOEM.0000040153.26156.ed>
- Capdeferro, N. & Romero, M. (2012). Are online learners frustrated with collaborative learning experiences? *The International Review of Research in Open and Distance Learning*, 13(2), 26–44.
- Cashen, V. M. & Leicht, K. L. (1970). Role of the isolation effect in a formal educational setting. *Journal of Educational Psychology*.
- Chang, H. H. & Chuang, S.-S. (2011). Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management*, 48(1), 9–18. <https://doi.org/10.1016/j.im.2010.11.001>
- Chen, Y.-H., Jang, S.-J., & Chen, P.-J. (2015). Using wikis and collaborative learning for science teachers' professional development. *Journal of Computer Assisted Learning*, 31(4), 330–344. <https://doi.org/10.1111/jcal.12095>
- Cheng, K.-H., Liang, J.-C., & Tsai, C.-C. (2013). The Role of Internet-Specific Epistemic Beliefs and Self-Regulation in High School Students' Online Academic Help Seeking: A Structural Equation Modeling Analysis. *Journal of Educational Computing Research*, 48(4), 469–489. <https://doi.org/10.2190/EC.48.4.d>
- Chi, M. T. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. *International handbook of research on conceptual change*, 61–82.
- Cole, M. (2009). Using Wiki technology to support student engagement: Lessons from the trenches. *Computers & Education*, 52(1), 141–146. <https://doi.org/10.1016/j.compedu.2008.07.003>

- Cress, U. & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning*, 3(2), 105–122. <https://10.1007/s11412-007-9035-z>
- Cunningham, W. (2002). Wiki: What Is Wiki. Retrieved from <http://www.wiki.org/wiki.cgi?WhatIsWiki>
- Daxenberger, J., Ferschke, O., Gurevych, I., & Zesch, T. (2014). DKPro TC: A Java-based Framework for Supervised Learning Experiments on Textual Data. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. System Demonstrations*.
- De La Paz, S., Swanson, P. N., & Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *Journal of Educational Psychology*, 90(3), 448–460. <https://10.1037/0022-0663.90.3.448>
- de Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 87–103. <https://10.1007/s11412-007-9006-4>
- Deci, E. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 43–70). Hillsdale, NJ: Erlbaum.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer.
- Deci, E. L. & Ryan, R. M. (2003). Intrinsic motivation inventory. *Self-Determination Theory*, 267.

-
- Delacre, M., Lakens, D., & Leys, C. (2017). Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test. *International Review of Social Psychology, 30*(1), 92. <https://10.5334/irsp.82>
- Dijksterhuis, A., van Knippenberg, A., Kruglanski, A. W., & Schaper, C. (1996). Motivated Social Cognition: Need for Closure Effects on Memory and Judgment. *Journal of Experimental Social Psychology, 32*(3), 254–270. <https://10.1006/jesp.1996.0012>
- Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *Three worlds of CSCL. Can we support CSCL?*, 61–91.
- Dillenbourg, P. & Fischer, F. (2007). Computer-supported collaborative learning: The basics. *Zeitschrift für Berufs-und Wirtschaftspädagogik, 21*, 111–130.
- Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The Evolution of Research on Computer-Supported Collaborative Learning. In D. N. Balacheff, D. S. Ludvigsen, D. T. d. Jong, D. A. Lazonder, & D. S. Barnes (Eds.), *Technology-Enhanced Learning* (pp. 3–19). DOI: 10.1007/978-1-4020-9827-7_1. Springer Netherlands.
- Doise, W., Mugny, G., & Perret-Clermont, A.-N. (1975). Social interaction and the development of cognitive operations. *European Journal of Social Psychology, 5*(3), 367–383. <https://10.1002/ejsp.2420050309>
- Dreu, C. K. W. d., Koole, S. L., & Oldersma, F. L. (1999). On the Seizing and Freezing of Negotiator Inferences: Need for Cognitive Closure Moderates the Use of Heuristics in Negotiation. *Personality and Social Psychology Bulletin, 25*(3), 348–362. <https://10.1177/0146167299025003007>
- Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly, 44*(2), 350–383. <https://10.2307/2666999>

- Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, 1(1), 3–14. <https://10.1016/j.edurev.2005.11.001>
- Efklides, A. (2008). Metacognition - Defining Its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13(4), 277–287. <https://10.1027/1016-9040.13.4.277>
- Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, 25(4), 949–960. <https://10.1016/j.chb.2009.04.004>
- Erkens, G., Jaspers, J., Prangma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, 21(3), 463–486. <https://10.1016/j.chb.2004.10.038>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://10.3758/BF03193146>
- Feldstein, A. (2011). Deconstructing Wikipedia: Collaborative Content Creation in an Open Process Platform. *Procedia - Social and Behavioral Sciences*, 26, 76–84. <https://10.1016/j.sbspro.2011.10.564>
- Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist*, 48(1), 56–66. <https://10.1080/00461520.2012.748005>
- Flanagin, A. J. & Metzger, M. J. (2000). Perceptions of Internet Information Credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515–540. <https://10.1177/107769900007700304>
- Flavell, J. H. (1974). The development of inferences about others. In *Understanding other persons* (pp. xv, 266). Oxford, England: Rowman and Littlefield.

-
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive—developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fletcher, T. D. (2010). Psychometric: Applied psychometric theory. R package version 2.2. Retrieved from <https://CRAN.R-project.org/package=psychometric>
- Flöck, F., Laniado, D., Stadthaus, F., & Acosta, M. (2015). Towards Better Visual Tools for Exploring Wikipedia Article Development—The Use Case of “Gamergate Controversy”. In *Proceedings of the Workshop on Wikipedia, a Social Pedia: Research Challenges and Opportunities*, page (to appear), Oxford, UK. (Cited on page 148).
- Flower, L. & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>
- Ford, T. E. & Kruglanski, A. W. (1995). Effects of Epistemic Motivations on the Use of Accessible Constructs in Social Judgment. *Personality and Social Psychology Bulletin*, 21(9), 950–962. <https://doi.org/10.1177/0146167295219009>
- Fournier-Viger, P., Gomariz, A., Campos, M., & Thomas, R. (2014). Fast Vertical Mining of Sequential Patterns Using Co-occurrence Information. In V. S. Tseng, T. B. Ho, Z.-H. Zhou, A. L. P. Chen, & H.-Y. Kao (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 40–52). Springer International Publishing. https://doi.org/10.1007/978-3-319-06608-0_4
- Franklin, T. & Harmelen, M. v. (2007). *Web 2.0 for Learning and Teaching in Higher Education*. The Observatory of Borderless Higher Education. London.
- Giffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin*, 68(2), 104–120.

- Graham, S. (1997). Executive control in the revising of students with learning and writing difficulties. *Journal of Educational Psychology*, 89(2), 223–234. <https://doi.org/10.1037/0022-0663.89.2.223>
- Gronau, Q. F., Ly, A., & Wagenmakers, E.-J. (2017). Informed Bayesian t-Tests. *arXiv:1704.02479 [stat]*. arXiv: 1704.02479.
- Hauck, W. W. & Anderson, S. (1984). A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *Journal of Pharmacokinetics and Pharmacodynamics*, 12(1), 83–91.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Heimbuch, S. & Bodemer, D. (2014). Supporting awareness of content-related controversies in a wiki-based learning environment. In *Proceedings of The International Symposium on Open Collaboration* (30:1–4). New York, NY, USA: ACM. <https://doi.org/10.1145/2641580.2641607>
- Heimbuch, S. & Bodemer, D. (2015a). Let's Talk about Talks: Supporting Knowledge Exchange Processes on Wiki Discussion Pages. In *Ninth International AAAI Conference on Weblogs and Social Media (ICWSM-15)* (pp. 56–61). AAAI Press.
- Heimbuch, S. & Bodemer, D. (2015b). Need for cognitive closure as determinant for guidance in Wiki-based learning. In T. Matsui, A. F. Mohd Ayub, B. Jiang, H. Ogata, W. Chen, S. C. Kong, & F. Qiu (Eds.), *Proceedings of the 23rd International Conference on Computers in Education (ICCE 2015)* (Vol. WIPP, pp. 10–12). Hangzhou, China: Asia-Pacific Society for Computers in Education.
- Heimbuch, S. & Bodemer, D. (2016a). Effects of implicit guidance on contribution quality in a wiki-based learning environment. In *Proceedings of the 12th International Conference of the Learning Sciences (ICLS 2016)* (Vol. 2, pp. 906–909). International Society of the Learning Sciences.

-
- Heimbuch, S. & Bodemer, D. (2016b). Wiki Editors' Acceptance of Additional Guidance on Talk Pages. In *Tenth International AAAI Conference on Web and Social Media (ICWSM-16)* (pp. 51–52). AAAI Press.
- Heimbuch, S. & Bodemer, D. (2017). Controversy awareness on evidence-led discussions as guidance for students in wiki-based learning. *The Internet and Higher Education*, 33(2017), 1–14. <https://10.1016/j.iheduc.2016.12.001>
- Heimbuch, S., Ollesch, L., & Bodemer, D. (2016). Effects of collaboration scripts on learning activities in wikis. EARLI SIG 7: Learning and Instruction with Computers. Dijon, France: European Association for Research on Learning and Instruction.
- Heimbuch, S., Uhde, K., & Bodemer, D. (2014). "a new Wiki way?" - An experimental study of collaborative knowledge building scripts. In A. F. Mohd Ayub, T. Matsui, C.-C. Liu, H. Ogata, S. C. Kong, & A. Kashihara (Eds.), *Proceedings of the 22nd International Conference on Computers in Education (ICCE 2014)* (Vol. WIPP, pp. 10–12). Nara, Japan: Asia-Pacific Society for Computers in Education.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion; psychological studies of opinion change*. OCLC: 187639. New Haven: Yale University Press.
- Ioannou, A., Brown, S. W., & Artino, A. R. (2015). Wikis and forums for collaborative problem-based activity: A systematic comparison of learners' interactions. *The Internet and Higher Education*, 24, 35–45. <https://10.1016/j.iheduc.2014.09.001>
- Isaacs, E. A. & Clark, H. H. (1987). References in Conversation Between Experts and Novices. *Journal of Experimental Psychology*, 116(1), 26–37. <https://10.1037/0096-3445.116.1.26>

- Janssen, J. & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist*, 48(1), 40–55. <https://10.1080/00461520.2012.749153>
- Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science*, 38(1), 59–88. <https://10.1007/s11251-008-9078-1>
- Järvelä, S., Volet, S., & Hana Järvenoja. (2010). Research on Motivation in Collaborative Learning: Moving Beyond the Cognitive–Situative Divide and Combining Individual and Social Processes. *Educational Psychologist*, 45(1), 15–27. <https://10.1080/00461520903433539>
- Johnson, D. W. & Johnson, R. T. (1985). Classroom conflict: Controversy versus debate in learning groups. *American Educational Research Journal*, 22(2), 237–256. <https://10.3102/00028312022002237>
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (2000). Constructive controversy: The educative power of intellectual conflict. *Change: The Magazine of Higher Learning*, 32(1), 28–37. <https://10.1080/00091380009602706>
- Johnson, D. W., Johnson, R. T., & Tjosvold, D. (2000). Constructive controversy: The value of intellectual opposition. In M. Deutsch P. T. Coleman (Ed.), *The handbook of conflict resolution: Theory and practice* (pp. 65–85). San Francisco, CA: Jossey-Bass.
- Kasemvilas, S. & Olfman, L. (2009). Design alternatives for a MediaWiki to support collaborative writing. *Journal of Information, Information Technology, and Organizations*, 4, 87–106.
- Keller, J. M. (1987a, September 1). Development and use of the ARCS model of instructional design. *Journal of instructional development*, 10(3), 2. <https://10.1007/BF02905780>

-
- Keller, J. M. (1987b, October 1). Strategies for stimulating the motivation to learn. *Performance + Instruction*, 26(8), 1–7. <https://10.1002/pfi.4160260802>
- Kelley, K. (2017). MBESS: The MBESS r package. R package version 4.4.0. Retrieved from <https://CRAN.R-project.org/package=MBESS>
- Kent, C., Laslo, E., & Rafaeli, S. (2016). Interactivity in online discussions and learning outcomes. *Computers & Education*, 97, 116–128. <https://10.1016/j.compedu.2016.03.002>
- Kimmerle, J., Moskaliuk, J., & Cress, U. (2011). Using wikis for learning and knowledge building: Results of an experimental study. *Educational Technology & Society*, 14(4), 138–148.
- Kimmerle, J., Moskaliuk, J., Oeberst, A., & Cress, U. (2015). Learning and collective knowledge construction with social media: A process-oriented perspective. *Educational Psychologist*, 50(2), 120–137. <https://10.1080/00461520.2015.1036273>
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. https://10.1207/s15326985ep4102_1
- Kittur, A. & Kraut, R. E. (2008). Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (pp. 37–46). New York, NY, USA: ACM. <https://10.1145/1460563.1460572>
- Kittur, A. & Kraut, R. E. (2010). Beyond Wikipedia- coordination and conflict in online production groups. (pp. 215–224). ACM Press. <https://10.1145/1718918.1718959>

- Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts – a conceptual analysis. *Educational Psychology Review*, 18(2), 159–185. <https://10.1007/s10648-006-9007-2>
- Konstan, J. A. & Riedl, J. (2003). Collaborative Filtering: Supporting Social Navigation in Large, Crowded Infospaces. In K. H. MSc PhLic, D. Benyon, & A. J. Munro (Eds.), *Designing Information Spaces: The Social Navigation Approach* (pp. 43–82). Springer London.
- Koo, D.-M. & Choi, Y.-Y. (2010). Knowledge search and people with high epistemic curiosity. *Computers in Human Behavior*, 26(1), 12–22. <https://10.1016/j.chb.2009.08.013>
- Kruglanski, A. W. (1990). Motivations for judging and knowing: Implications for causal attribution. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* (Vol. 2, pp. 333–368). New York, NY, USA: Guilford Press.
- Kruglanski, A. W. & Mayseless, O. (1987). Motivational effects in the social comparison of opinions. *Journal of Personality and Social Psychology*, 53(5), 834–842. <https://10.1037/0022-3514.53.5.834>
- Kruglanski, A. W. & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, 103(2), 263–283. <https://10.1037/0033-295X.103.2.263>
- Kruglanski, A. W. & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19(5), 448–468. [https://10.1016/0022-1031\(83\)90022-7](https://10.1016/0022-1031(83)90022-7)
- Kuckartz, U. (2014, January 23). *Qualitative text analysis: A guide to methods, practice and using software*. Google-Books-ID: 9B2VAgAAQBAJ. SAGE.

-
- Kuhn, D. (2000). Metacognitive Development. *Current Directions in Psychological Science*, 9(5), 178–181. <https://10.1111/1467-8721.00088>
- Kump, B., Moskaliuk, J., Dennerlein, S., & Ley, T. (2013). Tracing knowledge co-evolution in a realistic course setting: A wiki-based field experiment. *Computers & Education*, 69, 60–70. <https://10.1016/j.compedu.2013.06.015>
- Lai, Y. C. & Ng, E. M. (2011). Using wikis to develop student teachers' learning, teaching, and assessment capabilities. *The Internet and Higher Education*, 14(1), 15–26. <https://10.1016/j.iheduc.2010.06.001>
- Lakens, D. (2017a). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*.
- Lakens, D. (2017b). TOSTER: Two one-sided tests (TOST) equivalence testing. R package version 0.2.5. Retrieved from <https://CRAN.R-project.org/package=TOSTER>
- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger (Ed.), *HCI and Usability for Education and Work* (5298, pp. 63–76). Lecture Notes in Computer Science. DOI: 10.1007/978-3-540-89350-9_6. Springer Berlin Heidelberg.
- Lee, M. D. & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. DOI: 10.1017/CBO9781139087759. Cambridge: Cambridge University Press. Retrieved September 29, 2017, from <http://ebooks.cambridge.org/ref/id/CBO9781139087759>
- Leppink, J. (2010). Adjusting cognitive load to the student's level of expertise for increasing motivation to learn. In *Proceedings of the Eighth International Conference on Teaching Statistics [Conference paper]*. Slovenia: Ljubljana.
- Leppink, J., van Gog, T., Paas, F., & Sweller, J. (2015). Cognitive load theory: Researching and planning teaching to maximise learning. In J. Cleland & S. J.

- Durning (Eds.), *Researching Medical Education* (1st, pp. 207–218). Chichester, UK: John Wiley & Sons Ltd.
- Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology, 88*(4), 587–604. <https://doi.org/10.1037/0021-9010.88.4.587>
- Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences, 44*(7), 1585–1595. <https://doi.org/10.1016/j.paid.2008.01.014>
- Litman, J. A., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition & Emotion, 19*(4), 559–582. <https://doi.org/10.1080/02699930441000427>
- Litman, J. A. & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diverse and specific components. *Journal of Personality Assessment, 80*(1), 75–86. https://doi.org/10.1207/S15327752JPA8001_16
- Lops, P., Gemmis, M. d., & Semeraro, G. (2011). Content-based Recommender Systems: State of the Art and Trends. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 73–105). Springer US. https://doi.org/10.1007/978-0-387-85820-3_3
- Lowry, N. & Johnson, D. W. (1981). Effects of Controversy on Epistemic Curiosity, Achievement, and Attitudes. *The Journal of Social Psychology, 115*(1), 31–43. <https://doi.org/10.1080/00224545.1981.9711985>
- Lucassen, T. & Schraagen, J. M. (2011). Evaluating WikiTrust: A trust support tool for Wikipedia. *First Monday, 16*(5). <https://doi.org/10.5210/fm.v16i5.3070>
- Lucassen, T. & Schraagen, J. M. (2012). Propensity to trust and the influence of source and medium cues in credibility evaluation. *Journal of Information Science, 38*(6), 566–577. <https://doi.org/10.1177/0165551512459921>

-
- Luhmann, N. (1984). *Soziale Systeme: Grundriss einer allgemeinen Theorie*. Suhrkamp.
- Luhmann, N. (1997). *Die Gesellschaft der Gesellschaft*. Suhrkamp.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19–40. <https://10.1037/1082-989X.7.1.19>
- Makara, K. A. & Karabenick, S. A. (2013). Characterizing sources of academic help in the age of expanding educational technology: A new conceptual framework. *Advances in help seeking research and applications: The role of information and communication technologies*, 37–72.
- Marinho, L. B., Nanopoulos, A., Schmidt-Thieme, L., Jäschke, R., Hotho, A., Stumme, G., & Symeonidis, P. (2011). Social Tagging Recommender Systems. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 615–644). Springer US. https://10.1007/978-0-387-85820-3_19
- Matschke, C., Moskaliuk, J., & Kimmerle, J. (2013). The impact of group membership on collaborative learning with wikis. *Cyberpsychology, Behavior and Social Networking*, 16(2), 127–131. <https://10.1089/cyber.2012.0254>
- Maxwell, S. E. & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, 113(1), 181–190. <https://10.1037/0033-2909.113.1.181>
- Mayordomo, R. M. & Onrubia, J. (2015). Work coordination and collaborative knowledge construction in a small group collaborative virtual task. *The Internet and Higher Education*, 25, 96–104. <https://10.1016/j.iheduc.2015.02.003>
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Klagenfurt.

- McCroskey, J. C. & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs*, 66(1), 90–103. <https://doi.org/10.1080/03637759909376464>
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication*, 60(3), 413–439. <https://doi.org/10.1111/j.1460-2466.2010.01488.x>
- Mock, A. (2017). Open (ed) Classroom—Who cares? *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 28, 57–65.
- Morey, R. D. & Rouder, J. N. (2015). BayesFactor: Computation of bayes factors for common designs. R package version 0.9.12-2. Retrieved from <https://CRAN.R-project.org/package=BayesFactor>
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2009). Wiki-supported learning and knowledge building: Effects of incongruity between knowledge and information. *Journal of Computer Assisted Learning*, 25(6), 549–561. <https://doi.org/10.1111/j.1365-2729.2009.00331.x>
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2012). Collaborative knowledge building with wikis: The impact of redundancy and polarity. *Computers & Education*, 58(4), 1049–1057. <https://doi.org/10.1016/j.compedu.2011.11.024>
- Mugny, G., Butera, F., Sanchez-Mazas, M., & Perez, J. A. (1995). Judgments in conflict: The conflict elaboration theory of social influence. In B. Boothe, R. Hirsig, A. Helminger, & R. Volkart (Eds.), *Perception, evaluation, interpretation*. Seattle, WA: Hogrefe & Huber.
- Mugny, G. & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, 8(2), 181–192. <https://doi.org/10.1002/ejsp.2420080204>

-
- Nakamura, J. & Csikszentmihalyi, M. (2014). The Concept of Flow. In *Flow and the Foundations of Positive Psychology* (pp. 239–263). DOI: 10.1007/978-94-017-9088-8_16. Springer Netherlands.
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of learning & motivation* (Vol. 26, pp. 125–173). New York, NY: Academic Press.
- Neuenhaus, Nora. (2011). *Metakognition und Leistung: Eine Längsschnittuntersuchung in den Bereichen Lesen und Englisch bei Schülerinnen und Schülern der fünften und sechsten Jahrgangsstufe*. (Doctoral dissertation, Otto - Friedrich - Universität Bamberg, Bamberg).
- Noroozi, O., Biemans, H. J., Weinberger, A., Mulder, M., & Chizari, M. (2013). Scripting for construction of a transactive memory system in multidisciplinary CSCL environments. *Learning and Instruction*, 25, 1–12. <https://doi.org/10.1016/j.learninstruc.2012.10.002>
- Notari, M., Reynolds, R. B., Chu, S., & Honegger, B. D. (2016). *The wiki way of learning: Creating learning experiences using collaborative web pages*. OCLC: 942600516.
- Nov, O. (2007). What motivates Wikipedians? *Communications of the ACM*, 50(11), 60–64. <https://doi.org/10.1145/1297797.1297798>
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, 6(2), 257–268. <https://doi.org/10.1037/1528-3542.6.2.257>
- Nurmela, K., Lehtinen, E., & Palonen, T. (2000). Evaluating CSCL Log Files by Social Network Analysis. In C. Hoadley & J. Roschelle (Eds.), *Computer support for collaborative learning (CSCL) 1999 : Designing new media for a new millenium: Collaborative technology for learning, education and training, December 12-15, 1999, Stanford University, Palo Alto, California*. S.l.: Erlbaum.

- Papadopoulos, P., Demetriadis, S., & Weinberger, A. (2013). 'Make it explicit!': Improving collaboration through increase of script coercion. *Journal of Computer Assisted Learning*, 29(4), 383–398. <https://10.1111/jcal.12014>
- Piaget, J. (1977a). Problems of equilibration. In M. H. Appel & L. S. Goldberg (Eds.), *Topics in Cognitive Development* (pp. 3–13). Springer US.
- Piaget, J. (1977b). *The development of thought: Equilibration of cognitive structures*. Viking Press.
- Posner, I. R. & Baecker, R. M. (1992). How people write together. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, 1992* (Vol. 4, pp. 127–138). <https://10.1109/HICSS.1992.183420>
- Rad, H. S. & Barbosa, D. (2012). Identifying controversial articles in Wikipedia: A comparative study. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration* (7:1–7:10). New York, NY: ACM. <https://10.1145/2462932.2462942>
- Reiterer, S., Stettinger, M., Jeran, M., Eixelsberger, W., & Wundara, M. (2015). Advantages of Extending Wiki Pages with Knowledge-based Recommendations. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business* (46:1–46:4). New York, NY, USA: ACM. <https://10.1145/2809563.2809611>
- Renner, B. (2006). Curiosity about people: The development of a social curiosity measure in adults. *Journal of Personality Assessment*, 87(3), 305–316. https://10.1207/s15327752jpa8703_11
- Richardson, W. (2006). *Blogs, Wikis, Podcasts, and Other Powerful Web Tools for Classrooms*. Corwin Press.

-
- Rienties, B., Tempelaar, D., Van den Bossche, P., Gijsselaers, W., & Segers, M. (2009). The role of academic motivation in Computer-Supported Collaborative Learning. *Computers in Human Behavior*, 25(6), 1195–1206. <https://10.1016/j.chb.2009.05.012>
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney u test. *Behavioral Ecology*, 17(4), 688–690. <https://10.1093/beheco/ark016>
- Scardamalia und Bereiter. (2006). Knowledge Building: Theory, Pedagogy, and Technology. In *Cambridge Handbook of the Learning Sciences* (1st ed., pp. 97–118). New York: Cambridge University Press.
- Scardamalia, M. & Bereiter, C. (1983). The development of evaluative, diagnostic and remedial capabilities in children's composing. In M. Martlew (Ed.), (pp. 67–95). Wiley and Sons.
- Scardamalia, M. & Bereiter, C. (1985). Development of Dialectical processes in composition. In D. Olson, N. Torrance, & A. Hildyard (Eds.), (pp. 307–329). Cambridge University Press.
- Scardamalia, M. & Bereiter, C. (1986). Research on written Composition. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 778–803). New York, NY, USA: Macmillan.
- Scardamalia, M. & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, 3(3), 265–283. https://10.1207/s15327809jls0303_3
- Scardamalia, M. & Bereiter, C. (2003). Beyond Brainstorming: Sustained Creative Work with Ideas. *Education Canada*, 43(4), 4.

- Schlink, S. (2009). *Persönlichkeit entscheidet: Der Einfluss des Bedürfnisses nach kognitiver Geschlossenheit auf Entscheidungen unter Unsicherheit* (Doctoral dissertation, Universität Trier).
- Schlink, S. & Walther, E. (2007). Kurz und gut: Eine deutsche Kurzsкала zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit. *Zeitschrift für Sozialpsychologie*, 38(3), 153–161. <https://10.1024/0044-3514.38.3.153>
- Schönbrodt, F. D. & Wagenmakers, E.-J. (2017). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*. <https://10.3758/s13423-017-1230-y>
- Schuirman, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Pharmacodynamics*, 15(6), 657–680.
- So, H.-J., Seah, L. H., & Toh-Heng, H. L. (2010). Designing collaborative knowledge building environments accessible to all learners: Impacts and design challenges. *Computers & Education*, 54(2), 479–490. <https://10.1016/j.compedu.2009.08.031>
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge. Acting with Technology*. Cambridge MA: The MIT Press.
- Steiger, J. H. (2004). Beyond the f Test: Effect Size Confidence Intervals and Tests of Close Fit in the Analysis of Variance and Contrast Analysis. *Psychological Methods*, 9(2), 164–182. <https://10.1037/1082-989X.9.2.164>
- Suh, B., Chi, E. H., Pendleton, B., & Kittur, A. (2007). Us vs. Them: Understanding Social Dynamics in Wikipedia with Revert Graph Visualizations. In *IEEE Symposium on Visual Analytics Science and Technology, 2007. VAST 2007* (pp. 163–170). <https://10.1109/VAST.2007.4389010>

-
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. *Digital media, youth, and credibility*, 73100.
- Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. *Journal of Universal Computer Science*, 7(3), 254–277. <https://10.3217/jucs-007-03-0254>
- Suthers, D. D. (2003). Representational guidance for collaborative inquiry. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn* (pp. 27–46). Dordrecht, Netherlands: Springer.
- Tabachnick, B. G. & Fidell, L. S. (2012, June 14). *Using multivariate statistics* (6th ed.). Boston: Prentice Hall.
- The JASP Team. (2017). JASP: Jeffreys’s amazing statistics program. Computer Software version 0.8.2.0. Retrieved from <https://jasp-stats.org/>
- Trocky, N. M. & Buckley, K. M. (2016). Evaluating the impact of wikis on student learning outcomes: An integrative review. *Journal of Professional Nursing*. <https://10.1016/j.profnurs.2016.01.007>
- van Marle, H. J. F., Hermans, E. J., Qin, S., & Fernández, G. (2009). From Specificity to Sensitivity: How Acute Stress Affects Amygdala Processing of Biologically Salient Stimuli. *Biological Psychiatry*, 66(7), 649–655. <https://10.1016/j.biopsych.2009.05.014>
- Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1(1), 3–14. <https://10.1007/s11409-006-6893-0>
- von Glasersfeld, E. (1992). Questions and answers about radical constructivism. In M. Pearsall (Ed.), *Scope, sequence, and coordination of school science. volume II. relevant research* (pp. 169–182). Washington, DC: National Science Teachers Association.

- Vygotsky, L. S. (1986). *Thought and language (rev. ed.)* Cambridge, MA: MIT Press.
- Webster, D. M. & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), 1049–1062. <https://10.1037/0022-3514.67.6.1049>
- Webster, D. M., Richter, L., & Kruglanski, A. W. (1996). On Leaping to Conclusions When Feeling Tired: Mental Fatigue Effects on Impressional Primacy. *Journal of Experimental Social Psychology*, 32(2), 181–195. <https://10.1006/jesp.1996.0009>
- Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, 33(1), 1–30. <https://10.1007/s11251-004-2322-4>
- Weinberger, A. (2003). *Scripts for computer-supported collaborative learning. Effects of social and epistemic cooperation scripts on collaborative knowledge construction* (Text.PhDThesis, lmu).
- Weinberger, A., Stegmann, K., Fischer, F., & Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning* (pp. 191–211). New York, NY: Springer.
- Whitehead, J. L. (1968). Factors of source credibility. *Quarterly Journal of Speech*, 54(1), 59–63. <https://10.1080/00335636809382870>
- Wichmann, A. & Rummel, N. (2013). Improving revision in wiki-based writing: Coordination pays off. *Computers & Education*, 62, 262–270. <https://10.1016/j.compedu.2012.10.017>
- Wikipedia. (2016, December 14). *Wikipedia:neutral point of view*. In *Wikipedia*. Page Version ID: 754855206. Retrieved February 9, 2017, from https://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral_point_of_view&oldid=754855206

-
- Wikipedia. (2017a, July 14). *Wikipedia:BOLD, revert, discuss cycle*. In *Wikipedia*. Page Version ID: 790621044. Retrieved August 15, 2017, from https://en.wikipedia.org/w/index.php?title=Wikipedia:BOLD,_revert,_discuss_cycle&oldid=790621044
- Wikipedia. (2017b, January 25). *Wikipedia:featured content*. In *Wikipedia*. Page Version ID: 761887482. Retrieved February 9, 2017, from https://en.wikipedia.org/w/index.php?title=Portal:Featured_content&oldid=761887482
- Wikipedia. (2017c, January 22). *Wikipedia:statistics*. In *Wikipedia*. Page Version ID: 761282078. Retrieved February 9, 2017, from <https://en.wikipedia.org/w/index.php?title=Wikipedia:Statistics&oldid=761282078>
- Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. [Learning strategies of university students: Factor structure and reliability of a new questionnaire.] *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15(4), 185–200.
- Wilde et al. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, (15/2009), 31–45.
- Winter, S. & Krämer, N. C. (2014). A question of credibility – Effects of source cues and recommendations on information selection on news sites and blogs. *Communications*, 39(4), 435–456. <https://10.1515/commun-2014-0020>
- Winter, S., Metzger, M. J., & Flanagin, A. J. (2016). Selective Use of News Cues: A Multiple-Motive Perspective on Information Selection in Social Media Environments. *Journal of Communication*, 66(4), 669–693. <https://10.1111/jcom.12241>
- Zhang, X. & Zhu, F. (2006). Intrinsic motivation of open content contributors: The case of Wikipedia. In *Workshop on Information Systems and Economics*.

- Zheng, B., Niiya, M., & Warschauer, M. (2015). Wikis and collaborative learning in higher education. *Technology, Pedagogy and Education, 24*(3), 357–374. <https://doi.org/10.1080/1475939X.2014.948041>
- Zheng, Y. (2012). Unlocking founding team prior shared experience: A transactional memory system perspective. *Journal of Business Venturing, 27*(5), 577–591. <https://doi.org/10.1016/j.jbusvent.2011.11.001>
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2