



HAL
open science

Learning collaboratively with digital instructional media: Students' communicational behaviour and its influence on learning outcome

Daniel C. Heinrich, Mathias Hattermann, Alexander Salle, Stefanie Schumacher

► **To cite this version:**

Daniel C. Heinrich, Mathias Hattermann, Alexander Salle, Stefanie Schumacher. Learning collaboratively with digital instructional media: Students' communicational behaviour and its influence on learning outcome. INDRUM 2020, Université de Carthage, Université de Montpellier, Sep 2020, Cyberspace (virtually from Bizerte), Tunisia. <hal-03113846>

HAL Id: hal-03113846

<https://hal.science/hal-03113846v1>

Submitted on 18 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Learning collaboratively with digital instructional media: Students' communicational behaviour and its influence on learning outcome

Daniel C. Heinrich¹, Mathias Hattermann¹,
Alexander Salle², and Stefanie Schumacher²

¹Paderborn University, Germany, dch@math.upb.de

²Osnabrück University, Germany

In this paper, we present a theoretical instrument based on the interactive-constructive-active-passive framework to gauge the interactivity of dyads' communication processes in collaborative face-to-face learning scenarios. Subsequently, we show the applicability of the instrument to time-sampled video recordings of 63 pairs of students from different fields of study who learn descriptive statistics at tertiary level with different digital instructional media (e.g. video tutorials).

By relating the students' interactivity of their communication processes to their performance in pre- and post-tests we can show that a significant link between their communicational behaviour and the learning benefit exists across different kinds of digital learning materials and different fields of study.

Keywords: (Teachers' and) students' practices at university level, digital and other resources in university mathematics education, communication, collaborative learning, digital instructional media.

THE MAMDIM-PROJECT

During the last years, an increasing number of digital instructional media like video tutorials or commented presentations are used in university level mathematics courses across different fields of study (engineering, psychology, economics, ...). At the same time, a lack of research regarding the pedagogical design of digital learning resources and their collaborative aspects is stated (Borba, Askar, Engelbrecht, Gadanidis, Llinares, & Aguilar, 2016). To address this situation, the mamdim-project (*learning mathematics with digital media during the transition from secondary to tertiary education*) explores the usage and benefit of different digital instructional media. In cooperation with four German universities (University of Applied Sciences Pforzheim, Offenburg University of Applied Sciences, Bielefeld University and Brandenburg University of Technology) a total of almost 300 students of different fields of study learned with different instructional media dealing with descriptive statistics in different social forms (in dyads or as single learners). The data used in this research report is taken from the main study which took place in 2016 and 2017. Its design (pre-test | intervention | post-test) had been improved after a pilot study in 2015.

THEORETICAL BACKGROUND AND RESEARCH QUESTIONS

Spoken language and the associated communication processes play an important role in mathematics education research; an overview of the contemporary research can be

found in Morgan, Craig, Schuette & Wagner (2014). Steinbring (2015) explains that students' learning and understanding of mathematics is driven by interaction, which especially depends on the use of language and communication processes. Thus, it is not surprising that one of the main benefits of collaborative learning situations is the possibility to share ideas verbally with a learning partner. On the one hand, this advantage has been mentioned and the benefits of collaborative learning situations have been shown in many publications throughout the years (Dillenbourg, Baker, Blaye & O'Malley, 1996; Slavin, 1995).

On the other hand, a non-negligible number of studies show that collaborative learning in small groups does not lead to greater learning outcomes automatically (Barron, 2003). Indeed, our own study finds no significant differences between single learners and dyads in their post-test scores (for details, see Salle, Schumacher & Hattermann, 2019). Even a meta study by Lou, Abrami, Spence, Poulsen, Chambers, & d'Apollonia (1996) finds that in 28 % of published studies on that subject collaborative learning situations have zero or even a negative effect.

While these results might seem contradictory, they can be explained by use of the ICAP-framework by Chi and Wylie (2014) described in the following paragraph. For the origins of the framework, see Menekse, Stump, Krause & Chi (2013). Our instrument to assess communication within dyads is based on this framework.

Engagement activities, the CAP-hypothesis and the ICAP-framework

The ICAP-framework is designed to explain which types of interactions between learners are most effective in small group learning scenarios. To use this framework on a given situation, the overt learning activities of the students must be identified. Those observable interactions between learners or between a learner and the learning material are called *engagement activities* (Chi and Menekse, 2015). These are categorised and rank ordered with respect to their benefit to the students' learning outcome into *passive*, *active* and *constructive* activities (ibid). To give an example in terms of verbal communication, *passive* engagement describes a student who only listens to his learning partner or gives one-syllable responses (like "hm" or "okay"). A student communicating *actively* would be characterised by reading out loud from the given instruction material or repeating what has already been said by her/his learning partner. *Constructive* use of verbal communication will occur if the student elaborates on the learning material, tries to explain what is being said on her/his own or if she or he poses a question (Chi and Wylie, 2014). The CAP-hypothesis claims that with respect to the learning outcome, constructive behaviour is superior to active behaviour which in turn is more beneficial for learning than passive behaviour of learners. For an overview of theoretical considerations and empirical studies that support this hypothesis, see Chi and Wylie (2014). By means of this framework the seemingly contradictory positions in research can be explained. To give an example, a constructively behaving student will not be able to profit from a collaborative situation if the learning partner behaves only passively.

While these categories can be used in single learner scenarios, *only* within collaborative situations a fourth category, namely *interactive engagement*, is introduced. In this category, both partners need to behave *and* communicate constructively in the sense explained above and a sufficient degree of back-and-forth utterances between the students must occur (Chi and Wylie, 2014). For example, if a student raises a question on a specific topic and the other student responds to this question, the dyad behaves interactively. The two constructively learning students will then, in theory, profit from their collaboration because of their interactivity.

This background and the long-term aim to foster students' learning outcome in digital media learning scenarios at tertiary level leads to the following research questions: Can the interactivity of a dyad's face-to-face communication while learning mathematics with a digital instructional medium be linked to their learning outcome in a pre-post-test-scenario? Is this relation applicable for students of different fields of study and across different types of digital instructional media?

STUDY DESIGN AND TEST INSTRUMENT

To answer this question, video recordings of 63 pairs of first year students from different universities and fields of studies have been analysed:

- 42 engineering and economics students, University of Applied Sciences Pforzheim
- 40 psychology students, Bielefeld University
- 44 students becoming primary school teachers, Bielefeld University

During the intervention phase of our study, which lasted at most 70 minutes, each of the three groups used a different instructional medium to learn descriptive statistics (i. e. measures of central tendency and spread). In Pforzheim, a digital script in a moodle-environment containing definitions, formulas, examples, explanations but no further multimedia elements (like audio commentary or video clips) were present (cf. Figure 1). As an interactive element, short multiple-choice questions were incorporated into the material. As an example, following the slides dealing with the harmonic mean of a data set (cf. Figure 1) the following multiple-choice question is stated:

A garden centre creates a substrate by blending same masses of four different soils. These soils are known to have the following densities: Soil A: 710 kg/m^3 ; Soil B: 920 kg/m^3 ; Soil C: 830 kg/m^3 and Soil D: 1000 kg/m^3 . Calculate the average density of the substrate.

Possible answers were 851 kg/m^3 , 865 kg/m^3 and 857 kg/m^3 with the first one being the correct answer (calculated using the harmonic mean). The second (wrong) answer is calculated using the arithmetic mean of the four values and acts as a distractor. After answering the question, students got a direct feedback whether their calculation was correct or not. At Bielefeld University, the psychology students worked with a PowerPoint presentation. Each slide contains formulas and explanations of the topic at hand. Further details are given to the students through an audio narration that can be stopped and rewound at any time. The teacher students at Bielefeld University used a

series of animated video tutorials with audio narration to learn descriptive statistics. Each topic is addressed in its own segment. After each segment, the videos are paused automatically giving students time for discussion of the topic at hand before continuing to the next topic. For example, the video explaining the harmonic mean has a running time of three minutes and thirteen seconds (cf. Figure 2).

Especially the latter two instructional media include acoustic multimedia elements (audio narration) which could influence dyads' dialoguing behaviour compared to classical pen-and-paper scenarios. In this regard the moodle-environment used at Pforzheim University can be seen as a benchmark of our methodology because its design is very similar to a classic textbook and lacks audio elements.

Measures of central tendency

The harmonic mean

The harmonic mean of n numbers x_1, x_2, \dots, x_n is defined by

$$\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

The question arises in which situation the harmonic mean has to be used to calculate the average. The answer is given in the following.

If the values in question x_i are given as ratios, e.g. speeds (= length/time), it is crucial whether the data in the data set is referring to the length (nominator) or time (denominator). If the data refer to the time, i.e. the denominator, the average speed is calculated using the arithmetic mean; but if the data refer to the length, i.e. the nominator, the harmonic mean has to be used.

This is illustrated in the following example:

Figure 1: Slide from the instructional material used in Pforzheim dealing with the harmonic mean (translation by the authors).

Lisa's Performance in the Final

	1st Attempt	2nd Attempt	3rd Attempt
Lisa's Time	2:42	2:36	4:00
Quotient Words per Minute	74	77	50

Attempt →

Time for 200 Words →

Average found on Ranking →

64.52

Total Number of Attempts →

$$\bar{x}_{\text{harm}} = \frac{3}{\frac{1}{74} + \frac{1}{77} + \frac{1}{50}} = 64.52$$

Reciprocal of 2nd Attempt

Harmonic Mean

Used to calculate the average of rates, if they have a fixed nominator.

Figure 2: Frame from Bielefeld University (Teacher Students) instructional material dealing with the harmonic mean (translation by the authors).

The media intervention phases of all students have been videotaped and their computer screens were captured. Before and after each media intervention all students had to take the same pre- and post-tests consisting of both multiple-choice items and open questions regarding descriptive statistics. A detailed overview of the test items used can be found in Salle, Schumacher & Hattermann (2019).

Methodology

Based on the ICAP-framework described above and on methods derived from Chi and Menekse (2015), we developed a theoretical instrument to measure a dyad's interactivity based on video recordings, which is explained in the following; see also Hattermann, Heinrich, Salle & Schumacher (2018).

Measuring interactivity: the dialogue pattern score

To analyse the video recordings, the time-sampling method by Bakeman and Gottman (1997) was used. Each video has been organised into intervals of 10 seconds each and those intervals in which students communicate (verbally) for more than 5 seconds about the mathematical topic at hand were identified. This task was carried out by four qualified persons. To establish intercoder reliability, 13 % of the data set was coded by all four persons leading to a Krippendorff's alpha of 0.77 which represents a satisfying value (cf. Lombard, Snyder-Duch & Bracken, 2002).

Based on the ICAP-framework, the interactivity between the learners in each 10-second-segment is then rated on an ordinal scale ranging from 1 to 3 according to the coding scheme illustrated in Table 1 (see Chi and Menekse, 2015; Hattermann et al., 2018).

score	description
	Student communication in this 10 second segment ...
1	... is dominated by one student. (<i>active-passive or constructive-passive</i>)
2	... is performed by both partners, but not interactively. (<i>active-active, constructive-active, constructive-constructive</i>)
3	... has two constructive partners contributing interactively to it. (<i>constructive-constructive and interactive</i>)

Table 1: Coding scheme of communication interactivity.

As an example, the following dialogue is taken from a pair of psychology students working on the slide dealing with the harmonic mean (cf. Figure 2).

S1: One would need ... The arithmetic mean would be 200 characters because one would add $100 + 100 + 400$ and then one would divide by 3. That would be 200. That's the result of 600 divided by 3.

S2: Mhm.

S1: That's 200. This would be the arithmetic mean. But this is wrong because in that case he would have had to type faster.

S2: Mhm.

Student 1 tries to understand the given example of the harmonic mean and elaborates on the material given and thus behaves constructively. His partner, however, responds in one syllable mumblings, which is a passive activity. In terms of their interactivity, the corresponding segments in the video recording have to be coded with score 1.

This analysis has been carried out on all 63 recordings by two coders. To establish intercoder reliability, 7 % of the data set has been coded by both of them yielding a satisfying Krippendorff's alpha value of 0.81.

The interactivity of each dyad is then quantified by the *dialogue pattern score* (*dps*, Chi and Menekse, 2015; Hattermann et al., 2018) calculated as follows: The number of times each score (1, 2 or 3) occurs is counted for every dyad and the dialogue pattern score is calculated by taking the weighted average of the occurrences: For example, suppose the communication of a dyad consists of 19 coded segments in total. 7 of those segments were *score 1* segments, 8 of them reached *score 2* and the remaining 4 of them got a *score* of 3. From this, we can calculate the dialogue pattern score (*dps*) as

$$\frac{7 \cdot 1 + 8 \cdot 2 + 4 \cdot 3}{19} \approx 1.84.$$

Therefore, the communication of dyads with a dialogue pattern score close to 1 is dominated by single student contributions without interaction between the partners while scores closer to 3 represent a higher level of verbal interaction.

Measuring learning benefit: the normalised gain score

To link interactivity of dyads' communications with their learning outcome, we use the *normalised gain score* g (Hake, 1998) as a measure of learning benefit. This measure relates the pre- (x_{pre}) and post-test (x_{post}) results (as percentages) for each dyad by using the following formula:

$$g := \frac{x_{post} - x_{pre}}{1 - x_{pre}}$$

The post- and pre-test results are averages of both partners' results as percentage figures. This number relates the percentage points a learner actually gained between pre- and post-test to the percentage points he/she could have gained. For example, a student scoring 25 % in the pre-test and 50 % in the post-test achieved a normalised gain score of $g = 0.33$ (he/she gained 25 percentage points out of 75 percentage points he/she could have gained).

RESULTS

From the complete data set of 63 dyads, seven had to be excluded from further analysis since a lack of communication between the students made it impossible to calculate a

reliable dialogue pattern score. For the remaining 56 dyads, we gained information on the interactivity of the particular dyad (as measured by the *dps*) and the learning outcome (as measured by the normalised gain score *g*). To investigate whether dyads with a higher interactivity score on average show better learning outcomes than less interactive dyads, they are split into two categories: dyads with a high *dps* and dyads with a low *dps*. This is done by calculating the median dialogue pattern score for each of the three cohorts (Pforzheim, Bielefeld (Psychology) and Bielefeld (Teacher Students)) - dyads scoring not more than the median are considered “*low dps*” and dyads with a dialogue pattern score above the median are considered “*high dps*”. This analysis has been done for each group of students individually because the cohorts do not form a homogenous group as a whole: For example, they differ in their respective fields of study and the digital instructional media used during intervention and so the distribution of their dialogue pattern scores might vary.

Having constructed groups of *high dps* and *low dps* dyads within all three cohorts, the average normalised gain score for each group is calculated. All results are present within the following chart (Figure 3):

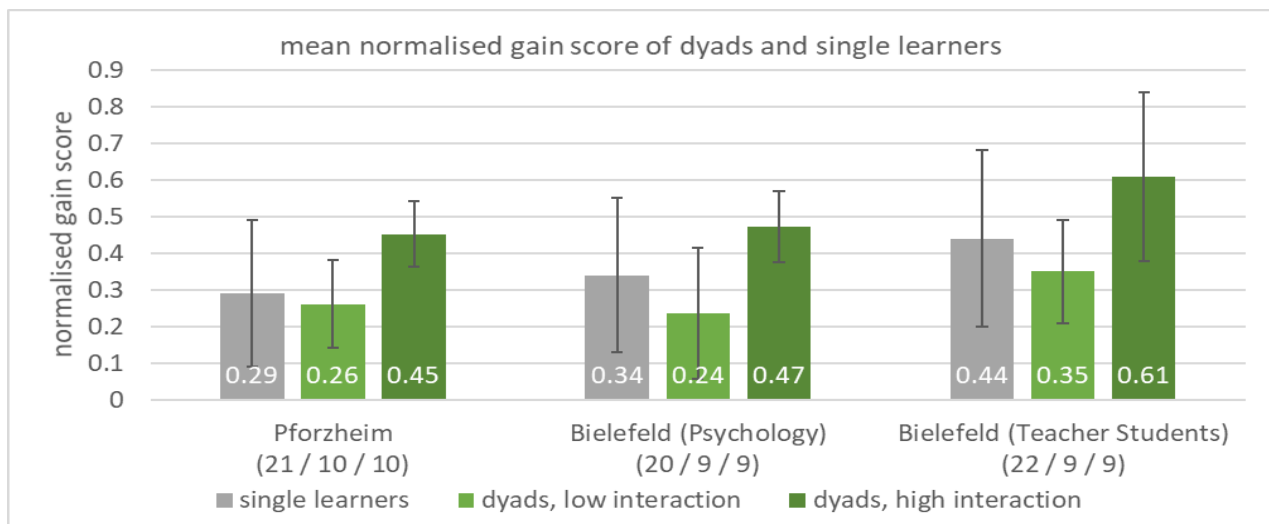


Figure 3: Average normalised gain scores with standard deviation for each group.

Independent t-tests carried out for each cohort show that the difference in means is highly significant for Pforzheim ($t = - 3.981$, $p = 0.001$) and Bielefeld (Psychology, $t = - 3.507$, $p = 0.004$) and significant for Bielefeld (Teacher Students, $t = - 2.886$, $p = 0.013$). This confirms a link between the dialogue pattern and the normalised gain score across different kinds of digital instructional media as well as students of different fields of study.

	Pforzheim	Bielefeld (Psychology)	Bielefeld (Teacher Students)
Correlation	0.595**	0.675**	0.754**

Table 2: Spearman Ranked Correlation coefficients between normalised gain and dialogue pattern score. ** indicates a highly significant correlation of $p < 0.01$

In order to confirm these findings, we test the two variables *dialogue pattern score* (dps) and *normalised gain score* (g) for correlation using Spearman's ranked correlation coefficient (Spearman's rho) as a non-parametric test. Pearson's correlation coefficient is not applicable here since we cannot assume linearity between the two variables. The correlation coefficients (cf. Table 2) confirm a highly significant correlation ($p < 0.01$) between the two constructs across all three cohorts.

Previous knowledge as a possible factor of influence

It cannot be ruled out that the significant increase in learning benefit between low- and high-interaction dyads can in part be explained by the influence of other variables. One possible factor of influence that is often linked with high achievement is given by the previous knowledge of the students in question. Although the normalized gain score takes the pre-tests results into account when measuring learning benefit, it is possible that students who are high achievers in mathematics can communicate about mathematics more easily and in greater detail – which can lead to deeper interactions with their learning partners. In that scenario a better pre-test score could explain both a high dialogue pattern score and a high learning benefit.

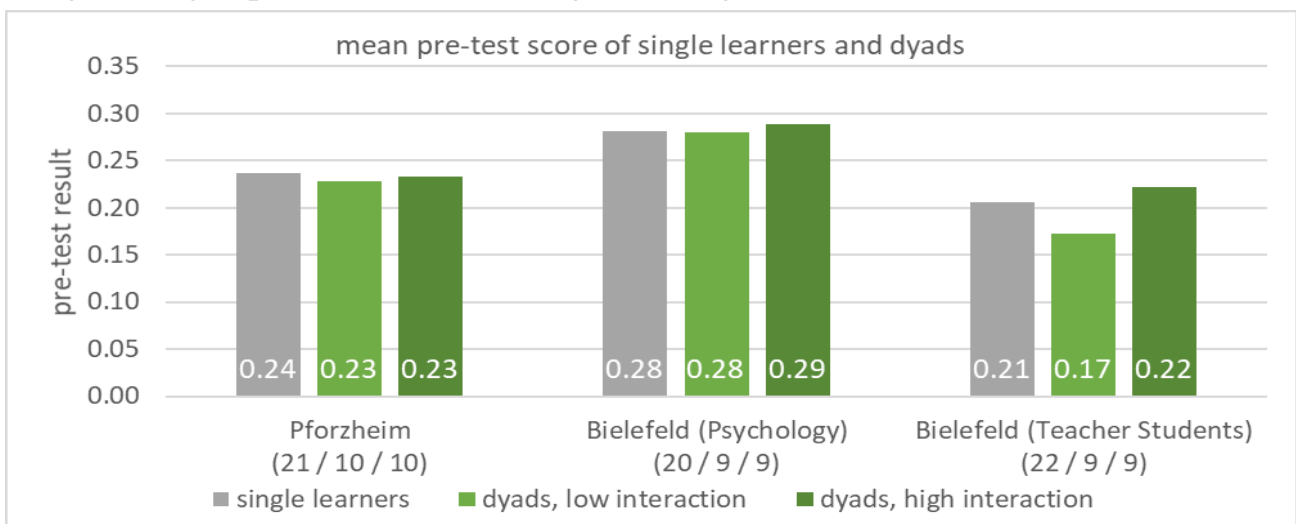


Figure 4: Comparison of mean pre-test results of high- and low-interaction dyads and single learners for each group.

To investigate this effect, we consider the groups of high- and low-interaction dyads as introduced in figure 3 above. For each of these groups we calculate and compare the mean pre-test scores – these can be found in figure 4. Regarding students in Pforzheim and the psychology students in Bielefeld, the mean pre-test scores between low- and high-interaction dyads are (nearly) identical (Pforzheim: $M_{low} = 0.23$, $M_{high} = 0.23$; Bielefeld (Psychology) $M_{low} = 0.28$, $M_{high} = 0.29$) while in the setting of the teacher students in Bielefeld a non-significant difference in means between those groups is present ($M_{low} = 0.17$, $M_{high} = 0.22$, $p = 0.338$).

As the previous knowledge of low- and high-interacting dyads in the domain of statistics is comparable in our study, it can be excluded as a main factor for explaining the higher learning benefit of high-interaction groups.

CONCLUSION AND PERSPECTIVES

Regarding our research questions, a significant link between the interactivity of the dyads' interaction, measured with the dialogue pattern score, and the learning outcome, measured by the normalised gain score, could be found using video recordings of 112 students in total. These results are in line with other research that has been done on this topic in different (non-digital) settings (Chi and Menekse, 2015). Furthermore, by using our framework to analyse students' communicational behaviour while learning with different types of digital instructional media, we were able to show that this link is not dependent on one specific type of media or a particular set of students. It remains an open question for future research whether the strength of this link is influenced by the type of digital instructional media used or the particular cohort of students considered. Indeed, the teacher students in Bielefeld show a higher overall learning benefit compared to other cohorts, but a slightly less significant result when comparing *high dps* and *low dps* dyads and more variance of the normalised gain score within these two subgroups. Furthermore, future research should take into account the possibility that the link between communication and learning benefit is dependent on the chosen mathematical content. To address this, future studies should focus on applying the ICAP framework on leaning materials dealing with domains outside of descriptive statistics.

Additionally, we are in the process to extend our analysis of possible factors of influence on both, the communicational behaviour and the learning outcome beyond the previous knowledge. For example, it is plausible that affective variables like students' interest or their domain specific self-efficacy could influence both, their interest to communicate with their learning partners and their learning benefit.

Using our method of analysing interactivity of dyads, we now aim to identify features of the digital instructional media which promote interaction between learners. This information could then be used to develop instructional material that fosters students' interactivity when learning collaboratively.

Acknowledgements: The project mamdim is funded by the German Federal Ministry of Education and Research BMBF (grant 01PB14011).

REFERENCES

- Bakeman, Roger; Gottman, John Mordechai (1997): Observing interaction. An introduction to sequential analysis. 2. ed. Cambridge [u.a.]: Cambridge University Press.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12, 307-359.
- Borba, M. C., Askar, P., Engelbrecht, J., Gadanidis, G., Llinares, S., & Aguilar, M. S. (2016). Blended learning, e-learning and mobile learning in mathematics education. *ZDM*, 48(5), 589–610

- Chi, M.T.H, & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4), 219-243.
- Chi, M.T.H, & Menekse, M. (2015). Dialogue patterns in peer collaboration that promote learning. In L. B. Resnick, C. Asterhan, & S, Clarke, (Eds.), *Socializing Intelligence Through Academic Talk and Dialogue* (pp. 263-274). American Educational Research Association.
- Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spada, & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189-211). Oxford: Elsevier.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64-74.
- Hattermann, M., Heinrich, D., Salle, A., & Schumacher, S. (2018): Instrument to analyse dyads' communication at tertiary level. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.). *Proc. 42nd Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 27-34). Umeå, Sweden: PME.
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content Analysis in Mass Communication. Assessment and Reporting of Intercoder Reliability. *Human Communication Research*, 28(4), 587–604.
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-Class Grouping: A Meta-Analysis. *Review of Educational Research*, 66(4), 423–458.
- Menekse, M., Stump, G., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, 102, 346-374.
- Morgan, C., Craig, T., Schuette, M., & Wagner, D. (2014). Language and communication in mathematics education: an overview of research in the field. *ZDM*, 46(6), 843-853.
- Salle, A., Schumacher, S., & Hattermann, M. (Eds.). (2020, in prep.). *Mathematiklernen mit digitalen Medien – Ergebnisse des mamdim-Projekts*. Berlin, Heidelberg: Springer.
- Slavin, R. E. (1995). *Co-operative learning: Theory, research, and practice*. 2. Ed. Boston: Allyn & Bacon.
- Steinbring, H. (2015). Mathematical interaction shaped by communication, epistemological constraints and enactivism. *ZDM*, 47(2), 281-293.