

jero

3
—
2020

*Journal for
Educational Research
Online*

Journal für Bildungsforschung Online

WAXMANN

JERO
Journal for Educational Research Online
Journal für Bildungsforschung Online

Executive Editors:

Nele McElvany, Institut für Schulentwicklungsforschung, Technische Universität Dortmund

Cornelia Gräsel, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

Editorial Office:

Rahim Schaufelberger, Institut für Schulentwicklungsforschung, Technische Universität Dortmund

Michael Rochnia, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

E-mail: editorialoffice@j-e-r-o.com

Editors:

Rolf Becker (Bern), Oliver Böhm-Kasper (Bielefeld), Wolfgang Böttcher (Münster), Martin Bensen (Münster), Wilfried Bos (Dortmund), Hartmut Ditton (München), Cornelia Gräsel (Wuppertal), Stephan Huber (Zug), Olaf Köller (Kiel), Harm Kuper (Berlin), Katharina Maag Merki (Zürich), Romain Martin (Luxemburg), Nele McElvany (Dortmund), Florian H. Müller (Klagenfurt), R. Daniel Muijs (Southampton), Jost Reinecke (Bielefeld), Kerstin Schneider (Wuppertal), Claudia Schuchart (Wuppertal), Wolfram Schulz (Camberwell), Birgit Spinath (Heidelberg), Petra Stanat (Berlin), Jan Van Damme (Leuven), Stefanie van Ophuysen (Münster), Andreas Voss (Hamburg), Horst Weishaupt (Frankfurt), Ludger Wößmann (München), Olga Zlatkin-Troitschanskaia (Mainz)

Please subscribe to the e-mail alert at www.waxmann.com/jero if you want to be informed about the publication of each new issue. There is a spring edition and an autumn edition each year. The subscription of the journal is free of charge.

Open access at www.waxmann.com/jero

ISSN 1866-6671

© Waxmann Verlag GmbH 2020
Steinfurter Straße 555, 48159 Münster, Germany
Internet: www.waxmann.com, E-mail: info@waxmann.com

Setting: Stoddart Satz- und Layoutservice, Münster
All rights reserved

Inhalt

| | |
|---|-----|
| <i>Marc Bienefeld, Oliver Böhm-Kasper & Christine Demmer</i> Extrem wichtig und doch vernachlässigt: Die geringe Präsenz von Replikationsstudien in der Erziehungswissenschaft | 3 |
| <i>Christoph Mischo, Katrin Wolstein, Sabrina Tietze & Svenja Peters</i> Professionelle Wahrnehmung bei Kita-Fachkräften: Erfassung, Generalisierbarkeit und Zusammenhänge | 23 |
| <i>Ina E. Rüber</i> Kontinuität und Wandel zivilgesellschaftlicher Partizipation im Erwachsenenalter: Eine Frage der Bildung und des Lernens? | 50 |
| <i>Ann-Sophie Grub, Antje Biermann & Roland Brünken</i> Prozessbasierte Erfassung professioneller Wahrnehmung der Klassenführung bei (angehenden) Lehrkräften: Ein systematisches Review | 75 |
| <i>Lars Burghardt, Anja Linberg, Simone Lehl, Kira Konrad-Ristau</i> Die Bedeutung frühkindlicher häuslicher und institutioneller Lernumwelten für frühe mathematische Kompetenzen..... | 103 |

Contents

Marc Bienefeld, Oliver Böhm-Kasper & Christine Demmer

Highly recommended and yet neglected:
The rarity of replication studies in educational science 3

Christoph Mischo, Katrin Wolstein, Sabrina Tietze & Svenja Peters

Professional vision of early childhood teachers:
Assessment, generalizability and relationships 23

Ina E. Rüber

Continuation and changes in civil
participation during adulthood:
A matter of education and learning? 50

Ann-Sophie Grub, Antje Biermann & Roland Brünken

Process-based measurement of professional vision of
(prospective) teachers in the field of classroom management:
A systematic review 75

Lars Burghardt, Anja Linberg, Simone Lehl, Kira Konrad-Ristau

The relevance of the early years home and institutional learning
environments for early mathematical competencies 103

Marc Bienefeld, Oliver Böhm-Kasper & Christine Demmer

Highly recommended and yet neglected: The rarity of replication studies in educational science

Abstract

Within the quantitative paradigm of social sciences, replication studies are considered of high scientific value but at the same time only a small number of actual replication studies are carried out and published. This disproportion has repeatedly been pointed out in the methodological discourse. However, while in some disciplines, e.g. psychology, there is now a growing interest in the topic, there has been no comparable development in educational science. Hence, three questions are worth considering: Can the relevance that methodology assigns to replication studies also be applied to educational science? What role do replication studies play for the current work of educational researchers? And what explanations can be explored for possible discrepancies between the designated relevance and the occurrence of replication studies? These questions determine the structure of this article: We will first discuss the methodological concept of replication studies and its relevance for educational research. In the second part of the article, the results of a qualitative explorative study will be presented. We carried out a total of 13 interviews with experts from various backgrounds in educational research, focusing on the above-mentioned questions. Beside some noteworthy similarities, the results show one main dividing line with regard to argumentation, which is the experts' personal affiliation with the quantitative or qualitative research paradigm respectively.

Keywords

Replication; Qualitative research methods; Quantitative research methods; Expert interviews; Evidence based research; Repetition

Marc Bienefeld, M.A. (corresponding author) · Prof. Dr. Oliver Böhm-Kasper · Jun.-Prof. Christine Demmer, Universität Bielefeld, Universitätsstraße 25, 33615 Bielefeld, Germany
e-mail: m.bienefeld@uni-bielefeld.de
oliver.boehm-kasper@uni-bielefeld.de
christine.demmer@uni-bielefeld.de

Extrem wichtig und doch vernachlässigt: Die geringe Präsenz von Replikationsstudien in der Erziehungswissenschaft

Zusammenfassung

*Vor allem innerhalb des quantitativen Paradigmas wird dem Replikationskonzept eine große Bedeutung zugeschrieben. Gleichzeitig werden aber nur wenige Replikationsstudien veröffentlicht. Auf dieses Missverhältnis wurde im methodischen Diskurs immer wieder hingewiesen. Während jedoch in einigen Disziplinen, wie z.B. der Psychologie, das Interesse an dem Thema wächst, gibt es in der Erziehungswissenschaft keine vergleichbare Entwicklung. Daraus ergeben sich folgende drei Forschungsfragen: Lässt sich die allgemeine theoretische Bedeutungszuschreibung auf die Erziehungswissenschaft übertragen? Welche praktische Relevanz haben Replikationsstudien im Alltag von Erziehungswissenschaftler*innen? Welche Erklärungen lassen sich für mögliche Diskrepanzen finden? Diese Fragen bilden die Grundlage des folgenden Artikels: Zunächst werden das theoretische Konzept von Replikationsstudien und deren Relevanz für die Erziehungswissenschaft diskutiert. Im zweiten Teil des Artikels werden die Ergebnisse einer qualitativ-explorativen Studie vorgestellt. Hierfür wurden 13 Interviews mit Expert*innen aus verschiedenen Funktionsbereichen der Erziehungswissenschaft bezüglich ihrer Einschätzung der zuvor vorgestellten Fragestellungen durchgeführt. Neben durchaus sichtbaren argumentativen Gemeinsamkeiten der Teilnehmer*innen erweist sich bei den Ergebnissen vor allem die persönliche Zuordnung der Expert*innen zum quantitativen bzw. qualitativen Forschungsparadigma als eine bedeutsame Trennlinie.*

Schlagworte

*Replikation; Qualitative Forschungsmethoden; Quantitative Forschungsmethoden; Expert*inneninterviews; Evidenzbasierte Forschung; Wiederholung*

1. Introduction

Quantitative-empirical research attributes great importance to the concept of replicating studies in order to provide a (more stringent) test of existing findings (replication) (see Popper, 1935; King, 1995; Schmidt, 2012). Not only in the natural sciences but also in other empirical sciences, replication emerges as “a very important method for testing empirical knowledge claims based on experimental and quantitative research” (Deutsche Forschungsgemeinschaft, 2017, p. 1). This demand for research results to be replicable methodologically is always emphasized in methodology textbooks (an example for pedagogical research is Newby, 2013). However, the concept of replication is implemented inadequately in research practice in the social sciences (see Schmidt, 2017). We can observe the radical mar-

ginalization of replication in pedagogical research in particular. A study by Makel and Plucker (2014) finds that only 0.13 % of the research results published in the 100 top-ranked journals of pedagogical research within a five-year period derived from replication studies. While there is a growing willingness to deal with the issue of replications (and associated methodological problems) in neighbouring disciplines such as psychology (see, for example Abele-Brehm et al., 2019; Camerer et al., 2018; Klein et al., 2018; Rost & Bienefeld, 2019), no similar development can currently be seen in educational science. The increasing attention that psychology is currently paying to the concept of replication largely follows the Open Science Collaboration (2015) study published in *Science Magazine* on the replicability of psychological research. This study finds that the statistically significant effects published in original studies can be reproduced in only about one third of matching replication studies. In addition, the study carries out a comparison of the mean effect sizes between original and replication studies. It finds that the mean effect size of the replication effects is only half the magnitude of the mean effect size of the original effects. The differences between the effect sizes of original and replication studies are significant here. Moreover, the authors test whether the effect sizes of the original studies fall within the 95 % confidence interval of the replication studies, which is the case in only about half of the studies analyzed (see Open Science Collaboration, 2015; controversial discussion in: Gilbert et al. 2016; Anderson et al., 2016). These findings provide the first empirical confirmation of the crisis of confidence in psychology that Pashler and Wagenmakers (2012) have postulated.

Despite the findings of Makel and Plucker (2014), a similar debate on replication has not yet been kindled in educational science. It is against this background that the present article aims to discuss the significance of the concept of replication for empirical research in educational science. We use the findings from an interview study to outline the value of the concept in the day-to-day research of educational scientists. For this purpose, we will first point out the methodological definitions and aims of replications (2.1) and relate them to educational science (2.2). We will present our interview study in which we asked research practitioners in educational science about their understanding of and experiences with replications: We will first explain the study design (3.1) and then present results from the interviews with quantitative and qualitative researchers (3.2).

2. The concept of replication and its relevance for educational science

2.1 Definitions and aims

Generally, replication refers to “the idea of reconfirming a scientific hypothesis or an experimental result by means of repetition” (Schmidt, 2012, p. 234). If we look for a universal and at the same time sufficiently nuanced definition that can serve

as the basis for actual research practice, then what emerges is more a “diffusion of concepts” (ibid., p. 235). There is a wide range of diverse and sometimes very contradictory approaches (see Cronbach, 1982; Sidman, 1960). Schmidt (2012) therefore proposes a “*functional approach*” (ibid., p. 240) that takes up the different aims of replication studies and uses them as the basis for classifying different forms of replication.

He proposes to first decide which elements should be kept constant in a replication study and which need to be altered and thereby provide a goal-oriented classification that can be transferred directly and explicitly to research practice

If, for example, the focus is on checking the sampling error, then the investigator should endeavour to keep all elements of the study constant, except for the sample. The recruitment strategy also remains the same; it is used again to obtain a new sample from the original population. If, on the other hand, the desired aim is to check lack of internal validity, then the application of the independent variable and the sample should remain constant. But as many alterations as possible should be made in the realization of the dependent variables. If certain findings should be tested for their validity for other populations, then only the population should be altered. Since all previously described aims are pursued on the basis of the same mode of investigation, such replications are referred to as “*direct replications*” (Schmidt, 2012, p. 244).

If in contrast the hypothesis is to be confirmed under completely different conditions, then the study design (e.g. sample selection, realization of dependent variables) itself has to be altered. The alterations are determined by the research question of the replication study. It is even possible to alter all elements except for the hypothesis which is then the only common basis of the two studies. This approach is referred to below as “*conceptual replication*” (see Schmidt, 2012).

Within research practice the functional approach can be seen as an ideal but hard to realize model of replication. Nevertheless, it provides a sufficiently subtle framework for replication studies to be conducted in a methodologically appropriate way, and it therefore forms the theoretical basis for the following sections.

2.2 To what extent can the concept of replication benefit educational science?

These statements show that the concept of replication is directed primarily at experimental studies and their replicability. In contrast to the natural sciences or psychology, however, actual experimental studies are scarce in educational research. This is due mainly to the structure of the field that educational science deals with: research questions in educational science usually address specific problems that occur under local conditions, and that therefore limit generalizations and the development of theories (see Berliner, 2002). He points out that research on education is generally determined by three factors that make replicability of research findings much more difficult: the power of context, the ubiquity of interac-

tions, and the dependence of research findings on processes of social and societal change (*decade by findings interactions*). Every educational activity takes place in a specific context: for example, family, kindergarten, school, or youth center. Each of these contexts has a unique structure, such as parenting behavior, municipal budgeting for pedagogical programs, leadership behavior of school heads, and how teachers teach. Due to the contextual influence of different local circumstances, study results are hard to compare. Even studies that have the same design may show different results. Here, context means above all integrating the human being in a complex and mutable network of social interactions. This becomes particularly clear if we think, for example, of teaching in a school class: the teacher has to adapt his or her actions to 20 to 30 unique combinations of student characteristics, such as intelligence, gender, eagerness and motivation to learn, socioeconomic background, and level of parental support. At the same time, student behavior is influenced by teacher characteristics, such as didactic and pedagogical competence, expectations in terms of what students can achieve, and even the well-being of the teacher at that moment. Given these complex interactions and their effect on pedagogically relevant features, the research field of educational science is much more heterogeneous than the artificial and highly controlled settings of laboratory experiments in the natural sciences and in psychology. Finally, the diverse interactions take place within their own specific social, cultural, and intellectual frameworks. Social and cultural change as well as progress in pedagogically relevant domains of knowledge (e.g. neurosciences) may cause a decade of research findings to be outdated or irrelevant already in the following decade (see Berliner, 2002).

Given this specific research field, it is worth discussing the question whether the concept of replication can be fruitful at all in educational science. We nonetheless believe that educational science should pay more attention to the concept of replication, and critically examine the robustness of its research findings. Our position is based on two lines of reasoning: first, educational science is regarded as a science of reality. Second, the idea of significance testing is central for the generalizability of empirical findings in educational science. As a science of reality, educational science seeks both to uncover regularities and to interpret an individual case embedded in specific local conditions (see Weber, 1904). However, given the complexity of the field we cannot adopt an either/or position with regard to these two aims, but instead embrace both. Educational research is characterized by such methodological diversity and by the high value placed on the discursive exchange of arguments (see Berliner, 2002). If educational researchers choose the quantitative research paradigm and apply statistical methods based on logical-deductive thinking, they inevitably face the question of whether their findings are robust and can be generalized. The above mentioned characteristics of the research field do not allow for a replication of all empirical findings. But this fact “must on no account be misappropriated as an excuse or apology for non-replicability in cases where the replicability of a scientific knowledge claim must be expected for methodological reasons” (see Deutsche Forschungsgemeinschaft, 2017, p. 3). If, as a science of reality, educational science takes seriously its aspiration to reveal also the

regularities of human actions and interactions, then we must not avoid a methodological debate on the value of replications.

Making empirical findings replicable is a demand that has always been woven into quantitative research in educational science: the idea of significance inevitably implies the idea of potential replicability, since the type I error (alpha error, i.e., the false rejection of a null hypothesis) may be minimized in its probability of occurrence, but can never be ruled out completely. Thus, a single study is not sufficient to decide whether the observation made was a clear finding or an artifact generated by error. It is at this point that the dilemma of non-experimental quantitative research becomes clear: on the one hand, the principle of traditional hypothesis testing inevitably implies the idea of the potential replicability of empirical observations, while on the other hand different contexts and complex interactions generate varying conditions for observations that cannot be replicated. One way out of this dilemma is the approach proposed by Schafer (2001): namely, to make a systematic link between replication studies and meta-analyses. Meta-analyses are not understood in this context as the statistical combination of more or less heterogeneous preparatory work. Instead, they are embedded in a research design in which replication studies are conducted in parallel in a heterogeneous research field. These replication studies serve as foundation for meta-analyses that create generalizable statements on mean effect sizes of pedagogical interventions or programs. Considering the challenges inherent in educational research (Berliner, 2002), this approach seems particularly useful and should be adopted more broadly in the discussion about a viable concept of replication.

3. Expert interviews on the value of replication studies in educational science

However, in order to enable this discussion within the discipline and beyond the purely methodological level, it seems unavoidable to deal with the following questions first: What value do educational researchers attribute to replication studies? And what reasons do these researchers give for the subordinate role of replication studies within educational research? However, as there are as yet hardly any reliable findings on this question, we decided on a qualitative exploratory approach to enable a first empirical address of this question.

3.1 Research design

Using expert interviews, we asked people who are involved in educational research in different central academic positions and who have insider knowledge with regard to our research question. Expert interviews are suitable for “reconstructing complex knowledge” (Meuser & Nagel, 2013) and for “acquiring practice-saturated

expert knowledge, the know-how of those who understand the laws and routines by which a social system reproduces, enacts and possibly modifies itself, or prevents this, but also the experiences of those who have designed and realized innovations” (ibid., pp. 457f.).

The interviewees were selected on the basis of four groups: editors of educational (research) journals, representatives of research funding, acknowledged researchers in the quantitative and the qualitative paradigm and in mixed methods, as well as researchers with experience in conducting replication studies. The latter group in particular are able to talk in an enlightening way about their experiences of the research process, which can be contrasted with the perhaps sceptical voices of other interviewees. Here, we follow the principle of a sampling that allows both minimal and maximal contrasts. In order to shed light on the conditions for and against replication studies, we take into account both individual motivations and levels of the academic system in the interviews. Of the 20 persons contacted, a total of 13 agreed to be interviewed¹ (see Table 1).

Table 1: Sample

| | Researchers | Research funding | Editors | Experience in replication studies |
|-----------------|-------------|------------------|---------|-----------------------------------|
| Intended sample | 8 | 4 | 4 | 4 |
| Realized sample | 8 | - | 3 | 2 |

Despite our repeated requests, we were unable to find a research funding representative willing to give an interview. As a result, this perspective and thus direct information are unfortunately not available. However, the interviewees were able to report indirectly on their experiences in obtaining third-party funding and on their activities as reviewers for research funding. In addition, our sampling technique of dividing the interviewees into four groups has not led to completely separate groups since some interviewees had double roles.

This applies above all to those who simultaneously embodied both the roles as editors and as empirical researchers. The interviewees clearly positioned themselves with regard to their own affiliation with the qualitative or quantitative research paradigm. This classification has turned out to be an important dividing line for the contrasting evaluation and therefore forms the basis for the presentation of our key findings.

All interviews were conducted by two persons, of whom at least one was a project manager of the study. The respective setting was chosen by the interviewees. The majority of interviews were conducted as face-to-face interviews in offices or

¹ For data privacy reasons, we are unable to provide a more explicit description of the sample. The description only relies on characteristics classified as relevant for the questions. This is the only way to ensure that no conclusions can be made about individual persons.

meeting rooms at the interviewees' workplaces (university or research institutes). Only one interview was conducted using a digital communication tool.

Following the elaborations of Gläser and Laudel (2009, p. 111) on the concept of the expert interview, the interviews were conducted with the help of a non-standardized guideline (see appendix). The guideline was compiled in accordance with the already mentioned theoretical and empirical publications on the topic of replication studies and their dissemination in social science (see 2.1). According to the questions asked, the main focus is on personal understanding, assessment of the significance of replication studies and previous contact with replication studies. In accordance with the principle of openness, the wording and order of the guideline as well as key topics in the interviews were adapted to the respective interviewees and interview situations (Gläser & Laudel, 2009, p. 150). This was particularly useful for the interviews with qualitative researchers: Since replications are not part of the qualitative research paradigm, we expected fundamental methodological questions that could only arise from the interview situation and that among other things depended on the respective qualitative research line represented by the interviewee.

The interviews were transcribed with the help of the transcription software f4. Since the focus in the analysis is on the thematic level, the transcription is presented as normal written German (Mayring 2016, p. 91).

The analysis was carried out with the help of the qualitative content analysis technique according to Glaser and Laudel (2009). In this technique, the categories used for the analysis are first derived deductively from the theoretical preliminary considerations. When in the course of the analysis text passages containing relevant information could not be integrated into the existing framework, the already existing categories were adapted or new categories were added inductively. In contrast to some classical approaches, however, no category derived from theory was removed. This ensured that the theoretical preliminary considerations were retained during the evaluation (see *ibid.*, p. 205). As already mentioned, the (self-) assignment of the interviewees to the qualitative or quantitative research paradigm has proved to be a relevant dividing line for the analysis. Therefore, this difference forms the basic structure for the following presentation of our findings:

We will first present the results from the interviews conducted with quantitative researchers, followed by the results from the interviews with qualitative researchers. These two parts differ in their structure. This is due to the fact that the interviews with the quantitative researchers closely followed the structure and contents of the guideline. The analysis of this group thus has a rather deductive character. The presentation of the results from the interviews with the quantitative researchers therefore closely follows the contents of the guideline (see 3.3.1). On the other hand, the interviews with the qualitative researchers already deviated from the structure of the guideline after the initial question, as we had anticipated. The qualitative researchers addressed aspects that had not been included in the theoretical guideline. In consequence, the presentation of the perspectives of the qualitative researchers on replication studies (see 3.3.2) is not based on the deductively

derived structure of the guideline, but on the inductively gained findings from the interviews.

3.2 Results from the interview study

3.2.1 Perspectives of quantitative researchers on replication studies

The interviewees whose work is mainly in the area of quantitative research methods attach as a matter of principle great importance to the theoretical concept of replication. They base their evaluation on the demand made in many methodology books for research results to be replicable. It should be noted, though, that they hold different understandings of the term *replication*.

While all interviewees mention the concept of direct replication (see 2.1), their comments differ greatly in terms of how broadly the concept can be understood. The range of possibilities that they mention extends from replications that are as similar as possible and that are expanded only by supplementary questions, to the investigation of a common hypothesis with the help of completely different research designs (Interview K, p. 1). What we notice terminologically is a recurring correspondence with the approach propounded by the American behavioral scientist Murray Sidman (1960), who distinguished between *direct* and *systematic* replications. However, none of the interviewees explicitly refers to this approach.

What they value in particular about replication studies is the fact that they can help researchers to protect the insights obtained from random influences and to generalize social mechanisms. Another potential task of replication studies in educational science is seen in the use of the secondary analysis of the original dataset to test methods of statistical analysis. In contrast to traditional secondary analyses, where already existing datasets are used for new research questions, the data are examined again in this context with regard to the same question, possibly using other methods of statistical analysis.

When it came to the question of how far the concept of replication is suitable for educational research, interviewees point out their specific research field, that is the relevance of context. While the description of this specific characteristic largely corresponds to what we have already explained in section 2.2, the interviewees nonetheless come to contrasting conclusions. Some state that replication studies are especially suited to dealing with possible changes to context. Others claim that the actual idea of replication studies is undermined here, and that for example long-term studies are a much better instrument for taking account of these conditions.

The interviewees in favor of replication studies also emphasize that the findings in educational science might not be of infinite duration but are nonetheless used to shape pedagogical practice and can therefore have far-reaching consequences. For

these interviewees, it is precisely replications that are the appropriate means to test once again whether results are still valid even after a longer period of time.

I'm also a bit amazed actually that it does not happen more often because I mean we all work with people. So that does not really differentiate us from medicine, and as soon as we do interventions, for example in schools, we should really say: so you are intervening in the system, you are manipulating the system in a particular way. And you have to be able to replicate; you have to show that what you are doing really has an effect. (C: 356)

And in this case replication studies are just simply worth their weight in gold, if you can just give suggestions with a greater degree of certainty. How you could shape reality. (A: 544)

3.2.1.1 Dissemination of replication studies and possible explanations

Despite their theoretical favor of replication studies, the interviewees have only had limited experience with replication studies in their actual research practice. This applies both to their own research practice as well as to their activities as editors or reviewers for research funding.

As a central reason for this, the interviewees mention the high level of pressure to innovate within the discipline, which means that they see working on areas that have not yet been researched as being more important and not least as more beneficial to their career than testing existing knowledge.

Because nobody wants to know: I get the same result that someone else has already found out before. So with us in the discipline, how I experience it, there's always a high level of pressure to throw new results onto the market and not to replicate. (C: 129)

According to the interviewees, this pressure is exacerbated by discipline-related third-party funding programs, which require both innovations and action frameworks as outcomes.

Although the interviewees occasionally report that results from replications submitted for publication were not published due to their lack of innovation, most do not think that there is a general policy of rejection among reviewers and editors when it comes to publishing replication studies. Rather, they report that such articles tend not to be submitted, and stress that they would even tend to appraise such articles positively – and precisely because they themselves attribute a high value to replication studies.

No, not submitted. We treat them like all the others, they go quite normally into (incomprehensible: review procedure? #00:01:30#) and I would even rather, I think, be a bit more benevolent, if something like that was done. (A: 37)

Besides the reasons already described, what the interviewees also mention as an obstacle is the research methodology of the discipline. They point out that experimental designs play hardly any role in educational research but they often play a crucial role for theoretical concepts of replication, or are even regarded as a necessary prerequisite. While some of the interviewees consider this a major obstacle for conducting such studies, others argue that replication studies should also be possible, for example, in standardized surveys since the goal is to generalize the findings made. For the interviewees, this generalization necessarily implies that results can be replicated, since otherwise a claim to generalization would inevitably be undermined. The statements made by the interviewees follow both our line of reasoning in section 2.2 on the implicit logic of replication of quantitative research in educational science, as well as the discussion of the functional approach in Schmidt (2012).

Representatives of this position also make the reservation, however, that for educational science it is above all notions of conceptual replication that are important. They put forth two questions for discussion: which conditions have to be met before a replication study can be categorized as a confirmation? Does educational science need its own concept of replication suited to its research practice? In addition, they point out that many studies in educational science have great methodological deficits that make replication almost impossible or that do not justify the effort required.

It becomes clear that the interviewees assume that the reasons they cite are interdependent and lead as a whole to the fact that a large number of those educational scientists working empirically see replications as uncreative and more of a hindrance to their career.

What is more, replications are always associated with a degree of scepticism towards existing research results. According to the interviewees, this methodologically justifiable scepticism is associated especially by junior researchers with the concern that they be regarded as “traitors” and block crucial career opportunities by testing or refuting published findings.

3.2.1.2 Opportunities to promote replication studies

If the aim is to raise the status that replications have in educational science, then according to the interviewees there are different ways of doing so.

Here, the interviewees consider it indispensable to raise awareness for methodological problems within educational science. They argue that only a stronger integration of the concept and value of replication studies in university courses, con-

ferences and publications can help to develop an understanding of which aims and benefits are associated with such studies. Moreover, a discipline-specific concept of replication should also be developed, one that properly addresses and takes into account the methodological reality of research in educational science. In order to promote this discussion, the interviewees consider it important that disciplinary associations and other academic organizations such as the Science Council participate and take a clear stance.

However, since university research today heavily relies on third-party funding, the theoretical debate can only be one component in the promotion of replication studies. Those providing research funding could help if they appraised projects by not only seeing innovation as being particularly important, but also by taking into account the value of replication studies, and if they ideally even established funding lines or an “award for replication studies” that specifically promote such projects.

Umm then / I think what is also needed is an impulse from those who fund research. That is those who distribute the money have to, so you can see that: as soon as a funding line is built up by politicians, for example, or by science policy, then it happens. So it also takes a bit of a push from the outside, because it's very time-consuming. (C: 297)

Reserved sections in journals and relevant calls for papers could also help promote replication studies.

The interviewees expect that these ways of promoting replication studies could enhance one another. For example, it seems entirely plausible that a larger number of specifically funded research projects will help to ensure that more such results are submitted and published. This, in turn, has the potential to show young researchers that replication studies are a key component in verifying results empirically, and do not stand in the way of an academic career.

3.2.2 Perspectives of qualitative researchers on replication studies

In qualitative research, replications are not a criterion of quality and are hence not discussed in methodology books. Nevertheless, if we include qualitative research in our study, then we are by no means implying that there is an unfulfilled methodological claim for such studies. Rather, it is a matter of exploring where qualitative research has potential links with and differences from the described functions that replications have for the logic of quantitative research. Thus, the line of questioning pursued by our project shifts when it comes to qualitative research:

- To what extent do qualitative researchers regard the replication of studies as a positive option (or not) for qualitative research?
- (Where) do qualitative researchers see elements in the qualitative research process that bear a connection with replication?

- (To what extent) do they think the replication of studies could open up interesting perspectives for qualitative research?

Unlike the more uniform premises of quantitative research, empirical approaches in qualitative research are characterized by a considerable variety of methodological principles. The interviewees are located in different schools of qualitative methodology, so that a certain range of research approaches such as ethnographic, as well as different reconstructive and interpretative approaches are represented in the sample². One interviewee had experience in conducting qualitative replications.

All interviewees assign the idea of replication in the narrower sense to the quantitative research paradigm and to approaches shaped by the natural sciences, such as medicine, physics, and also pedagogical-psychological research. What the interviewees identify as being typical of these approaches is a nomothetic research logic that aims for statistical representativeness and statements with general validity. Replications are seen here as a means to test the validity of results and research tools. This results in a clear demarcation line of qualitative-interpretative research from replications in the narrower sense. This demarcation is justified methodologically, but also – and this is particularly interesting for our research question – in relation to the educational research field (see 2.2). However, when they were presented with broader approaches of replication (see 2.1), some of the interviewees certainly saw worthwhile and potentially transferable aspects of replications for their own work:

Well, I would have said spontaneously: ‘This is a problem for quantitative researchers and I have absolutely nothing to do with it’. But that is not true ... [it’s] not as simple as that. (E: 12)

3.2.2.1 Methodological objections to replications from the perspective of qualitative research

The interviewees consider the principle of replication, whereby existing results are tested and checked, to be diametrically opposed to the fundamental premise regarding the openness of qualitative research. Thus, the research process is based on an essentially open-ended question. This research process may well comprise initial settings, but these are settings that act as theoretical heuristics and are intended to be open to dispute during the research process. Thus, the original question may shift again in line with the circular research process – something that is not compatible with the stringent testing of given results. While replications require that initial settings be kept constant, our interviewees believe that qualitative

² We did not interview researchers who place themselves methodologically within qualitative content analysis. Since qualitative content analysis is originally an approach of quantitative analysis, and quantifying logic is often used with the approach in qualitative projects, closer proximity to replication studies would be expected.

research uses dynamic concepts. Among other things, these concepts are constituted by the unique structure and logic of the respective research field and research object, which emerge successively. Assumptions that reduce complexity, as they are needed for variables in quantitative studies (e.g., in the form of a variable 'gender'), are therefore not set from the beginning in qualitative research. These presumptions are shifted into the research process itself in order to be able to open up complexity in the first place (e.g., "*What form does gender take here?*") (M: 26).

The interviewees also describe the value and handling of contextuality, as well as the relationship between subjectivity and standardization, in a very different way. Drawing on the theoretical supposition that research objects are historically, culturally and socially embedded, they argue that qualitative research is always concerned with concrete and complex cases. These cases are systematically considered in their contextual embedding under "*field conditions*" (D: 9), which opposes the idea of the replicability of results (*ibid.*). For the interviewees, influencing contexts should not be controlled and standardized in qualitative research, but rather firmly integrated. This is the complete opposite to experimental research, which is considered the classic area where replications are used. For in experiments, contextual factors as well as the subjectivity of researchers are kept as low or as constant as possible. In contrast, qualitative research usually incorporates natural elements of the data collection process. The interviewees conceive the data collection itself as an instance of social interaction that explicitly welcomes the unfolding of the researcher's subjective relevance. In order for this to emerge within research in the first place, this logic requires not as little but as much contextuality as possible, which is again considered contrary to replication: "*You have to have an idea of the high validity of context to want to replicate something*" (M: 29).

Unlike in quantitative approaches, the interviewees surmise that the knowledge generated is subject to "*probation*" (L: 14), but less so to rules and let alone to laws. Thus, according to the interviewees, there is in some strands of qualitative research the claim to generalization, but no claim to truth. Proven knowledge can become obsolete in a changing field and is only valid for the time being. It may therefore also make sense for qualitative research to investigate whether the knowledge generated is still valid. This may suggest a point of contact with the theoretical foundations of quantitative research and with replications. It entails, however, a new and fundamentally open, as well as theoretically grounded, research question – and less the pure testing of knowledge already gained (L: 16).

Thus, qualitative research is concerned less with making the scientific knowledge generated more reliable or with quantifying it, and more with exploring its complexities, such as when a particular research question is worked on anew: "*That's why I think that the basic figure is more difference than identity. And you would, I think, that is the idea of the research gap, do exactly what you can expect new knowledge from, and not what confirms existing knowledge*" (B: 46; M: 18).

The understanding of "*research gaps*" mentioned here is also cited by the quantitative researchers and points to a common understanding of science or a shared

scientific system in which the aim is to generate new knowledge and where there is sometimes reported a great pressure to innovate. However, the discovery of new connections and the generation of new theoretical concepts is part of the original self-understanding of many qualitative research approaches, so that this also touches on a core area of methodology.

Ethnographic researchers also emphasize that the value of methodology or of the approach underpinned by methodology varies within qualitative research approaches. Thus, the quality of an ethnographic study can only be guaranteed to a limited extent by means of a determinable methodological approach. Instead, reconstructive approaches such as the documentary method establish clearly structured evaluative steps and thereby demand a more standardized methodological approach.

3.2.2.2 The understanding of the research object and the research landscape of educational science

From the interviewees' point of view, the research objects of educational science cannot be studied separately from their respective concrete, complex and ever-changing contexts: "*We look at something different each time*" (L: 1). Therefore they are sceptical of replications aimed at testing the research tool and of mere adaptations of the debate on replications from other disciplines in which research objects appear less variable. Moreover, for the interviewees, other disciplines feature more uniform research approaches. In contrast, educational science is characterized by a distinctly differentiated research landscape that, with its heterogeneity of approaches, is not – nor has to be – universally compatible with the claim to replication. They clearly reject restricting educational research to 'evidence' because the concept taken from medicine and linked to replication does not do justice to the area of interest of educational science and to its heterogeneous research landscape.

3.2.2.3 The value of research for pedagogical practice and pedagogical professionalization

What may also be specific to the discipline is the argument that science and research do not directly provide recommendations for action because they do *not have the role of justifying practice, but rather of opening up possibilities for practice to reflect and gain a certain distance* (B: 42). Hence, research has the task of providing complex insights and concepts that enable pedagogical professionals to develop the ability to condense complexity into concrete decisions for action. Unlike in laboratory situations pedagogical practice is characterized by uncertainties and by influences that are hard to control. For that reason, there can in essence be no reliability of application that is, no guarantee that knowledge from educational science can be successfully put into practice. That also applies to rep-

licated knowledge and models. Thus, the “technology deficit” (Luhmann & Schorr, 1982) of pedagogy runs counter to the idea of a knowledge that is sufficiently verified through replication and that is therefore technologically applicable. According to the interviewees, disciplines in which replications have a high status are aimed more at adding value through practical application than at adding value through knowledge. For the interviewees, they are oriented less towards activity that prioritizes understanding and abductive reasoning, and more towards a “*what-works logic*” in practice.

3.2.2.4 Frictions in research policy and research psychology

The experts in qualitative research also say that replications in some ways run counter to the research system, which focuses primarily on innovation. A “*certain compulsion to originality*” is observed (E: 10). When it comes to a scholar’s own academic reputation and the success of his or her academic career, the interviewees also report considerable internal university pressure to generate third-party funding, and they find replications not helpful for this purpose. They also suspect that the publication policy of journals is not geared towards replication studies. Another hindrance are the political ambitions in which science is entangled. Especially for educational science they observe an unreflected actionism on the part of policy-makers who draw on research results and show little desire for lengthy and thorough proofs even in quantitative studies.

In addition, the qualitative researchers suspect that there is a “*state of truce*” (B: 26) among researching colleagues in which, in a figurative sense, everyone likes to “*tend their own garden*” (B: 12; also, J: 11). For the interviewees, the repetition of studies may also reveal potential inadequacies in the initial study that contradict the research claim, but nonetheless arise in practice (J: 11). Replications could therefore be considered a vote of no confidence, and could make a person seem like lacking loyalty amongst colleagues (D: 23). There is the suspicion that such a vote of no confidence is even more automatically directed at the researcher in qualitative research than in quantitative studies. This is because the entire research process and findings in qualitative research are more closely tied to the researcher, his or her methodological decisions and interpretations, while in quantitative research many research activities are standardized components of research tools (M: 34).

3.2.2.5 Potential analogies and links to conceptual replications in qualitative research?

Adopting a broad interpretation of the idea of replication, the interviewees also see similar mechanisms of testing or review in some aspects of qualitative research, for example in theoretical sampling. In this case, the sample is first drawn according to theoretical assumptions, but then successively and contrastively on the ba-

sis of initial results. Here, the following question is central: Are there similar cases that contradict the previous findings or could make them more nuanced? The sample is complete if other cases no longer bring new results to light, but only confirm the results already found. This step defines the study's scope of application, so that an analogy to the idea of replication can be seen (M: 14, E: 3). However, the interviewees also emphasize the factor of differentiation in this research step, which is directed more at investigating the variability within the phenomenon observed than at finding regularities across different contexts (D: 12).

The closest proximity to conceptual replication is attributed to research that aims at (theoretical) generalization (D: 7-8): "*Structural generalization could actually be open to replication*" (M: 9). In addition to theoretical sampling, additional factors that validate and therefore 'replicate' can be identified here. For example, in the procedural step of creating interpretations within objective hermeneutics (see Oevermann, 2000) (B: 36), where alternative meanings are developed in the discourse and examined along the case until one interpretation establishes itself and no interpretations that are more valid can be found (M: 9ff.). Since the case structure hypotheses thus formed can only claim validity until further notice, it would certainly be possible for another research group to 'test' these structure hypotheses using the same or indeed different material. However, the interviewees state that this only happens to a certain extent at conferences, where results are presented and discussed. This can be seen as a curtailed form of the 'plausibility check' (M: 5). At the same time, probation in academic discourse is seen as an important form of critical (self-)examination: 'validity' of results is measured in the reception and discussion of the results by other academics, by "*usefulness [for subsequent research], less by evidence*" (B: 9, 22, 37). In order to allow readers a critical methodological understanding and to take into account the quality criterion of transparency, the corresponding data material has to remain accessible (E: 6) – which, for example with regard to already archived material, immediately raises questions about anonymization, data privacy and research ethics (J: 11).

4. Summary and conclusion

In summary, a large proportion of those interviewees belonging to the quantitative paradigm attribute (in line with the methodological literature) a clear value to the concept of replication when it comes to safeguarding empirical research results. In contrast, those interviewees belonging to the qualitative paradigm see no direct links to replications in the narrower sense. Only one interviewee sees replications that have the aim of testing qualitative studies as being a potentially viable, but at the same time quite pointless, approach. On the other hand, the qualitative researchers deem the repetition of research questions and studies with deliberate variations (in particular, variations of sample, of historical-temporal context, of methodological or theoretical approach) as certainly being useful. They see the

aim of replication here as being to explore the socially or historically altered field of research, or to cast a new perspective on results in the light of current methodological and theoretical questions. However, in order to be able to deal with a certain research question in a comparative manner, e.g., a study presented forty years ago, it is necessary to adapt various facets of the research design, since both the research field and the methodological discussion have changed. While some of the interviewees see the concept of replication as an opportunity here, others argue for a clear conceptual demarcation to be made so as not to obscure the different methodological premises of quantitative and qualitative research approaches. Since pre-conceptions of replications are rather narrow, introducing replication into qualitative research contexts would lead to misconceptions about the aims and premises of qualitative research.

Despite attributing a very different value to replications, the researchers belonging to the quantitative paradigm also agree with the statement that the object of research in educational science is much more dependent on context than is the case in many disciplines in the natural sciences. However, this does not lead to the belief that replications cannot be applied in educational research and are of no benefit. Rather, it is precisely replication studies that may be a suitable means for enabling researchers to confirm or refute results under altered conditions. Due to the contextuality and other limitations of the research practise of educational science (e.g., limited research resources, changing staff, few experimental designs), the researchers in the quantitative paradigm also see direct replications as being barely feasible. On the other hand, most quantitative experts assign a significant value to the approach of conceptual replication.

The experts put forth some shared arguments for the low prevalence of replication studies. Thus, researchers from both research paradigms report of structural reasons that are in line with the DFG's position on the replicability of research results, which states: "The weight of quantitatively-parametrising control, evaluation and gratification systems prevailing in research today has the effect of creating increased (and continually increasing) pressure to compete and to accelerate results. This is manifested in decisions (and underlying decision criteria) about career moves, financial support, location of publication, and institutional structural trends" (Deutsche Forschungsgemeinschaft, 2017, p. 4). The interviewees speak in this context primarily of an enormous pressure within the discipline to publish and be innovative, a pressure that is increased not least by a system of research funding that places a high value on generating "new" knowledge, and that links support to a presentation of results that can be used directly for pedagogical practice. According to all interviews, this affects how replication studies are perceived within the discipline, with replication studies being deemed, and especially so by young researchers, as not being particularly career-enhancing.

According to the quantitative researchers, who largely accept the concept of replication, replication studies and their reputation in educational science should be specifically promoted in the future. What is required for this purpose is a deeper and broader theoretical discussion of the fundamental concept as well as the possi-

bilities of conducting replications (for example, in university teaching, and at conferences). But, an expansion of relevant third-party funding lines is also required. Such a process can be supported for example by editors of empirical journals that promote the publication of such results, or by academic associations that promote such studies.

This is different for qualitative research. Following the assessments of the experts, the specific functions of replications and the associated methodological challenges and consequences must be discussed first. It would be interesting to look at inherent understandings of ‘replications’ and theoretical generalization in the sense of typology or structure hypothesis within different qualitative approaches.

References

- Abele-Brehm, A. E., Gollwitzer, M., Steinberg, U., & Schönbrodt, F. (2019). Attitudes toward open science and public data sharing. A survey among members of the German psychological society. *Social Psychology* 50(4), 252–260.
- Anderson, C. J., Bahník, S., Barnezzt-Cowan, M., Bosco, F. A., Chandler, J., Chartier, C. R., & Zuni, K. (2016). Response to comment “Estimating the reproducibility of psychological science”. *Science*, 351, 1037-c. <https://doi.org/10.1126/science.aad9173>
- Berliner, D. C. (2002). Comment: Educational research: The hardest science of all. *Educational Researcher*, 31(8), 18–20.
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., & Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior*, 2, 637–644.
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Deutsche Forschungsgemeinschaft. (2017). *Replizierbarkeit von Forschungsergebnissen. Eine Stellungnahme der Deutschen Forschungsgemeinschaft*. Deutsche Forschungsgemeinschaft: Bonn.
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on „Estimating the reproducibility of psychological science“. *Science*, 351, 1037-b. <https://doi.org/10.1126/science.aad724>
- Gläser, J., & Laudel, G. (2009). *Experteninterviews und qualitative Inhaltsanalyse, als Instrumente rekonstruierender Untersuchungen* (3., überarbeitete Auflage). Wiesbaden: VS.
- Luhmann, N., & Schorr, K.E. (1982). Das Technologiedefizit der Erziehung und die Pädagogik. In N. Luhmann & K.E. Schorr (Eds.), *Zwischen Technologie und Selbstreferenz. Fragen an die Pädagogik* (pp. 11–41). Stuttgart: Suhrkamp.
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, 43(6), 304–316.
- Mayring, P. (2016). *Einführung in die qualitativ Sozialforschung: Eine Anleitung zu qualitativem Denken* (6. Auflage). Weinheim: Beltz.
- Meuser, M., & Nagel, U. (2013). Experteninterviews: Wissenssoziologische Voraussetzungen und methodische Durchführung. In B. Frieberthshäuser, A. Langer & A. Prengel (Eds.), *Handbuch qualitative Forschungsmethoden* (pp. 457–471). Weinheim: Beltz.
- King, G. (1995). Replication, replication. *Political Science and Politics*, 28(3), 444–452.
- Klein, R. A., Vianello, M., Hasselmann, F., Adams, B. G., Adams, R. B., Alper, S., & Nosek, B. A. (2018). Many labs 2: Investigating variation in replicability across

- samples and settings. *Advances in Methods and Practices in Psychological Science*, 1, 443–490.
- Newby, P. (2013). *Research methods for education*. Hoboken: Taylor and Francis.
- Oevermann, U. (2000). Die Methode der Fallrekonstruktion in der Grundlagenforschung sowie der klinischen und pädagogischen Praxis. In K. Krainer (Ed.), *Die Fallrekonstruktion* (pp. 58–157). Frankfurt a. M.: Suhrkamp.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251).
- Pashler, H., & Wagenmaker, E.-J. (2012). Editors introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, 7(6), 528–530.
- Popper, K. R. (1935). *Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft*. Wien: Springer.
- Rost D. H., & Bienefeld M. (2019). Nicht replizieren: publizieren!? *Zeitschrift für Pädagogische Psychologie*. 33(3-4), 163–176.
- Schafer, W. D. (2001). Replication: A design principle for field research. *Practical Assessment, Research & Evaluation*, 7(15).
- Schmidt, S. (2012). Muss man alles wiederholen? Eine kritische Analyse des Replikationsbegriffs in der modernen Wissenschaft. In W. Ambach (Ed.), *Experimentelle Psychologie in Grenzgebieten*, Vol. 10 (pp. 233–261). Würzburg: Ergon.
- Schmidt, S. (2017). Replication. In M. C. Makel & J. A. Plucker (Eds.), *Towards a more perfect psychology: Improving trust, accuracy, and transparency in research* (pp. 233–253). Washington, DC: American Psychological Association.
- Sidman, M. (1960). *Tactics of scientific research. Evaluating experimental data in psychology*. New York: Basic Book.
- Weber, M. (1904). Die “Objektivität” sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für Sozialwissenschaft und Sozialpolitik*, 19(1), 22–87.

Christoph Mischo, Katrin Wolstein, Sabrina Tietze & Svenja Peters

Professionelle Wahrnehmung bei Kita-Fachkräften: Erfassung, Generalisierbarkeit und Zusammenhänge

Zusammenfassung

Die Professionelle Wahrnehmung stellt einen wichtigen Kompetenzaspekt des pädagogischen Personals dar. Unter Professioneller Wahrnehmung wird die Fähigkeit verstanden, die lern- und entwicklungsrelevanten Tiefenmerkmale einer pädagogischen Interaktion a) wahrzunehmen bzw. zu identifizieren, b) zu bewerten, und c) daraufhin Handlungsalternativen zu generieren. Sie ist ein wissensgesteuerter Wahrnehmungsprozess, in den verdichtetes konzeptuelles Wissen und Erfahrungswissen einfließen. Im deutschen Sprachraum ist die Professionelle Wahrnehmung bei Fachkräften aus Kindertageseinrichtungen (Kitas) noch kaum untersucht. Die Studie Professionelle Wahrnehmung von Kita-Fachkräften (ProWaK) geht u. a. der Frage nach, ob die Professionelle Wahrnehmung bei Kita-Fachkräften mit Hilfe einer videobasierten Präsentation von vier typischen Interaktionssequenzen mit Hilfe der Methode des simultanen Lauten Denkens und retrospektiver Interviews erfasst werden kann. Inhaltlich werden dabei die Domänen der Emotionalen Unterstützung, der Organisation des Kita-Alltags und der Lernunterstützung unterschieden. Die Messung der Professionellen Wahrnehmung lässt sich daher als Ergebnis unterschiedlicher Messfacetten konzipieren: die Videosequenzen, die Befragungsmethode (simultanes Lautes Denken vs. retrospektives Interview), die erfassten Aspekte (Wahrnehmung/Identifikation, Interpretation/Bewertung, Handlungsalternativen) und die inhaltlichen Domänen. In einer Generalisierbarkeitsstudie wird überprüft, wie hoch der jeweilige Beitrag der Messfacetten ausfällt, und ob die Ergebnisse über die hier dargebotenen Videosequenzen hinaus generalisiert werden können. Zusammenhänge mit der Ausbildungsstufe und der Berufserfahrung deuten auf die Validität des Messverfahrens hin. Die Befunde werden in ihrer methodologischen und praktischen Bedeutung für die Professionalisierung des Personals in Kitas diskutiert.

Prof. Dr. Christoph Mischo (corresponding author) · Katrin Wolstein, M.Sc. · Sabrina Tietze, M.A. · Svenja Peters, M.A., Institut für Psychologie, Pädagogische Hochschule Freiburg, Kunzenweg 21, 79117 Freiburg, Deutschland
E-Mail: mischo@ph-freiburg.de
katrin.wolstein@ph-freiburg.de
sabrina.tietze@ph-freiburg.de
svenja.peters@ph-freiburg.de

Schlagworte

Professionelle Wahrnehmung; ErzieherIn; FrühpädagogIn; Lautes Denken; Generalisierbarkeitstheorie

Professional vision of early childhood teachers: Assessment, generalizability and relationships

Abstract

In recent years, the demands on staff in day care facilities have increased significantly. One aspect of professional competence refers to professional vision. Competency models and previous research describe professional vision as an important competence aspect of professionals in educational settings. Regarding professional vision, some researchers differentiate between the ability to a) perceive or identify relevant teaching events, b) interpret and evaluate these events and c) generate alternative options of action. Hence, professional vision is a knowledge-driven perception process which is affected by conceptual knowledge as well as experience-based knowledge. In German-speaking countries, professional vision of early childhood teachers has rarely been investigated. This study explores the question, whether professional vision can be assessed using a video-based presentation of interaction sequences using the thinking aloud-technique and a focused interview. By doing so, we differentiated the interaction domains of Emotional Support, Classroom Organization and Instructional Support. The contribution of each facet of measurement (way of assessment: thinking aloud vs. interview, video sequences, aspect of professional vision: perception/description, evaluation, generating options of action, and the domain of interaction) is investigated in a generalizability study. Furthermore, the generalizability of the presented video sequences is analyzed. Correlations with the educational level and work experience may be interpreted as an indication for the validity of the measurement. The findings are discussed in terms of methodological and practical issues regarding the field of early childhood professionals.

Keywords

Professional vision; Early childhood teacher; Thinking aloud-technique; Interview; Generalizability theory

1. Einleitung

Die Tätigkeitsbereiche für das Personal in Kindertageseinrichtungen (Kitas) sind vielfältig und umfassen die unmittelbare, primäre Arbeit mit dem Kind (Beziehungen gestalten, Bildungs- und Entwicklungsprozesse fördern) sowie die sekundäre Arbeit mit dem Umfeld (Arbeit mit Eltern und Bezugspersonen, institutions- und teambezogene Arbeit, Netzwerke gestalten; Pasternack &

Schulze, 2010). Die zentrale Aufgabe von Fachkräften in Kitas besteht darin, „die Selbstbildungspotenziale der Kinder durch eine angemessene und überlegte Gestaltung der Umwelt und durch das Aufgreifen und Erweitern der kindlichen Themen zu unterstützen und herauszufordern“ (Fröhlich-Gildhoff, Nentwig-Gesemann & Pietsch, 2011; Tietze & Viernickel, 2007, S. 13); dabei sind auch die Beziehungs- und Bindungsgestaltung notwendige Voraussetzungen (Pasternack & Schulze, 2010). So stellt eine sichere Bindung eine wichtige Voraussetzung für Explorationsverhalten und die Selbstbildung dar. Auch für das Lernen in sozialen Interaktionen (ko-konstruktivistisches Lernen) ist die soziale Beziehung und damit die Bindung eine wichtige Voraussetzung (Ahnert & Gappa, 2013; Lamb, 1998; Textor, 2007). Im beobachtbaren Verhalten zeigen die Fachkräfte in Deutschland den Aspekt der Bildung i. S. der Lernunterstützung jedoch in geringerem Ausmaß als die Beziehungs- und Bindungsgestaltung (Mackowiak, Wadepohl & Bossart, 2014; Tietze et al., 2012). Zwar liegen zahlreiche Befunde darüber vor, wie eine entwicklungsförderliche Interaktion der Fachkraft aussehen sollte (z. B. NICHD, 2002; Weltzien, 2014), die personenseitigen Voraussetzungen bei den Fachkräften sind jedoch weniger gut untersucht.

Über welche Kompetenzen die Fachkräfte verfügen sollen, wird in einschlägigen Kompetenzmodellen spezifiziert (Anders, 2012; Fröhlich-Gildhoff, Nentwig-Gesemann, Pietsch, Köhler & Koch, 2014). Neben den Komponenten des professionellen Wissens und professioneller Überzeugungen bzw. Einstellungen findet in der Frühpädagogik neuerdings die Wahrnehmung professioneller Situationen als eine Teilkompetenz Erwähnung (Fröhlich-Gildhoff et al., 2014). Diese Teilkompetenz ist in der Kindheitspädagogik – im Gegensatz zur Forschung über Lehrkräfte – in der anglo-amerikanischen Forschung erst in Ansätzen (Hamre et al., 2012), im deutschen Sprachraum dagegen noch kaum beforscht worden. Nach einer Skizze der Befunde zur Professionellen Wahrnehmung aus der Lehrerforschung und der Kindheitsforschung werden Erhebungsmethoden für die Professionelle Wahrnehmung dargestellt und herausgearbeitet, welche Konsequenzen daraus für eine reliable und generalisierbare Messung abgeleitet werden können, und welche biographischen Merkmale als Validitätsindikatoren in dieser Studie verwendet werden können.

2. Professionelle Wahrnehmung als wichtiges Kompetenzmerkmal

Die erfahrungs- und wissensbasierte Wahrnehmung, Bewertung und Interpretation beruflicher Situationen wird als Professionelle Wahrnehmung bezeichnet (van Es & Sherin, 2002). Vor allen Dingen bei Lehrkräften wurden im Rahmen des Expertenparadigmas Unterschiede zwischen ExpertInnen und NovizInnen in der Wahrnehmung von Unterrichtssituationen untersucht (Berliner, 2004). ExpertInnen verfügen über relevante Wahrnehmungskategorien (Clark & Peterson,

1986; Hogan & Rabinowitz, 2009; Peterson & Comeaux, 1987; Sherin & van Es, 2005), äußern mehr Kognitionen zu relevanten Aspekten des Lehr-Lerngeschehens (Krull, Oras & Siskask, 2007), sie weisen komplexere Wissensseinheiten auf (Chunks) und verbinden Situationswahrnehmungen mit Handlungsmöglichkeiten (Gobet, 2005). Auch in Deutschland wird die Professionelle Wahrnehmung von unterschiedlichen Gruppen von ForscherInnen intensiv untersucht (z. B. Blömeke, Kaiser & Lehmann, 2010; Gold, Hellermann & Holodynski, 2016; Seidel & Stürmer, 2014). Konzeptuell unterscheiden einige AutorInnen dabei die Fähigkeit zur *Wahrnehmung* bzw. Identifikation relevanter Unterrichtsereignisse („noticing“) von der *Interpretation* bzw. Bewertung dieser Ereignisse und die Generierung von Handlungsalternativen („reasoning“; van Es & Sherin, 2002; König et al., 2014; Seidel & Stürmer, 2014). Im konkreten Vollzug gehen die Prozesse der Wahrnehmung und der Interpretation bzw. Bewertung ineinander über (König et al., 2014, S. 78f.) und dürften eher als ein Kontinuum zu konzipieren sein (Blömeke, Gustafsson & Shavelson, 2015). Die Professionelle Wahrnehmung ist dabei nicht nur ein durch die objektiven Merkmale der Situation gesteuerter Prozess (bottom up), sondern wird ebenso durch die Erwartungen und kognitiven Kategorien des Wahrnehmenden beeinflusst (Top-down-Prozess). Daher ist eher von einer Verschränkung von Bottom-up- und Top-down-Prozessen auszugehen (Schweer, Thies & Lachner, 2016, S. 129ff.). Die Professionelle Wahrnehmung von Lehrkräften steht auch im Zusammenhang mit dem Lernerfolg der SchülerInnen (Kersting, Givvin, Thompson, Santagata & Stigler, 2012). Insbesondere in der Komponente der Interpretation und des Schlussfolgerns von Unterrichtssequenzen (reasoning) manifestiert sich die Qualität des deklarativen Professionellen Wissens von ExpertInnen, die in der Lage sind, ihre integrierten und vernetzten Wissensstrukturen situationsspezifisch anzuwenden (Berliner, 1987; Seidel & Stürmer, 2014; König et al., 2014).

Bei ErzieherInnen bzw. Fachkräften in Kitas steht die Professionelle Wahrnehmung dagegen deutlich weniger im Fokus der Forschungsbemühungen. In einer kleineren Studie mit ErzieherInnen in Deutschland konnte Faas (2013) zeigen, dass erfahrene Fachkräfte ein elaboriertes Praxis- und Theoriewissen bei der Interpretation beruflicher Episoden anwenden, die als Video oder Fotografie dargeboten wurden. In einer Studie der NICHD-Gruppe in den USA (Hamre et al., 2012) wurde gezeigt, dass eine videounterstützte, auf dem Classroom Assessment Scoring System (CLASS; Pianta, LaParo & Hamre, 2008) basierende Trainingsmaßnahme bei 440 ErzieherInnen die Wahrnehmung bzw. Identifikation („noticing“) qualitativ guter Fachkraft-Kind-Interaktionen verbessert, und dass dieser Fähigkeit eine Mediatorfunktion für die Vorhersage der eigenen Interaktionskompetenz der Fachkraft zukommt (hier: Lernunterstützung; Hamre et al., 2012). Der konzeptuelle Rahmen des CLASS-Instruments, in der Kita-Version Pre-K, besteht im Kern aus den drei Domänen der *Emotionalen Unterstützung* (Emotional Support), der *Organisation des Kita-Alltags* (Classroom Organization) und der *Lernunterstützung* (Instructional Support; Übersetzung nach Weltzien et al., 2017, S. 17). Die Einschätzung für die Domäne Emotionale Unterstützung er-

folgt in den Dimensionen Positives Klima, Negatives Klima, Feinfühligkeit und Orientierung am Kind. Die Domäne Organisation des Kita-Alltags differenziert sich in die Komponenten Verhaltensmanagement, Beschäftigungsgrad der Kinder und Lernarrangement. Die Domäne der Lernunterstützung bezieht sich auf die drei Dimensionen Kognitive Anregung, Feedbackqualität und Unterstützung sprachlichen Lernens. Die Interaktionsqualität wird durch zertifizierte BeobachterInnen in drei bis vier Zyklen à 15-20 Minuten für domänenspezifische Unterdimensionen auf einer siebenstufigen Skala eingeschätzt (1-2 = niedrige Qualität, 3-5 = mittlere Qualität, 6-7 = hohe Qualität). Aussagen über die Qualität der drei Domänen werden letztendlich über den Mittelwert der zugehörigen Dimensionswerte getroffen.

Die mit der CLASS erfasste Interaktionsqualität zeigte sich in groß angelegten Studien als prädiktiv für unterschiedliche Maße der kindlichen Entwicklung und des Lernens (soziale Kompetenz, geringere Verhaltensprobleme, Sprach- und Lesekompetenz sowie mathematische Kompetenz; z.B. Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Hamre & Pianta, 2005; Wasik & Hindman, 2011).

Eine auf der CLASS basierende Trainingsmaßnahme zur Professionellen Wahrnehmung findet als Video Assessment of Interactions and Learning (VAIL) in der Ausbildung von Fachkräften in den USA Anwendung (Jamil, Sabol, Hamre & Pianta, 2015). Ergebnisse einer groß angelegten Studie legen nahe, dass Verhaltensänderungen, Professionelles Wissen und die Professionelle Wahrnehmung transaktional verschränkt sind, und dass nicht nur Wissen und Professionelle Wahrnehmung Verhaltensänderungen vorhersagen, sondern dass auch Verhaltensänderungen zu einer Veränderung im Professionellen Wissen und zu einer Verbesserung der Professionellen Wahrnehmung führen (Pianta et al., 2014). Somit liegen auch im frühpädagogischen Bereich v. a. internationale Befunde vor, die für die Relevanz der Professionellen Wahrnehmung von Kita-Fachkräften als wichtiges Kompetenzmerkmal sprechen.

2.1 Erfassung der Professionellen Wahrnehmung

Grundsätzlich hat sich bei der Erfassung von situationsbezogenen und stärker verdichteten Schemata und Skripten die bild- und v. a. videounterstützte Präsentation von beruflichen Sequenzen bewährt (Bromme, 1992, S. 124; König et al., 2014; Sherin & van Es, 2005). Die videounterstützte Präsentation beruflicher Sequenzen dient dazu, bei den ProbandInnen Prozesse der Professionellen Wahrnehmung zu evozieren, da diese die Videosequenzen beschreiben, interpretieren und analysieren sollen.

Die Erfassung der Kognitionen bei der Wahrnehmung und Interpretation präsentierter Sequenzen erfolgt teilweise in einem offenen Antwortformat (z. B. Faas, 2013; Kersting; 2008), teilweise mit Hilfe von Rating-Items (z. B. Seidel & Stürmer, 2014). Das offene Antwortformat besteht dabei meist in einer Verschriftlichung der Kognitionen (Paper Pencil z. B. bei Schäfer & Seidel, 2015; online z. B. bei Hamre et al., 2012; Kersting, 2008). Wenngleich bestimmte Zusammenhänge zwischen

einem gebundenen und einem offenen Antwortformat bestehen (Schäfer & Seidel, 2015), so kann das offene Format doch als „state of the art“ bezeichnet werden (Schäfer & Seidel, 2015, S. 36).

Zur Erfassung simultaner Kognitionen bei Problemlöseaufgaben schlagen Ericsson und Simon (1993) die Methode des Lauten Denkens vor. Nach Ericsson (2006) weist diese Methode eine geringe Reaktivität und eine hohe Validität auf. Die Validität *retrospektiver* Beschreibungen und Interpretationen als Indikatoren *simultaner* und tatsächlich ablaufender kognitiver Prozesse sieht Ericsson (2006, S. 230f.) dagegen kritisch. Diese Kritik trifft vor allem auf die Verschriftlichung kognitiver Prozesse zu, da hierbei in besonderem Maße reflexive Prozesse beteiligt sind. Auch in einem anderen Bereich, nämlich bei der Erfassung von Lernstrategien, hat sich die Methode des Lauten Denkens besonders bewährt (Bannert, 2004).

Allerdings können auch mit der Methode des Lauten Denkens nur solche kognitiven Prozesse erfasst werden, die den Probanden bewusst sind. Die Unterscheidung zwischen bewusstem deklarativen Wissen und nicht bewusstem Wissen (tacit knowledge) geht u. a. auf Polanyi (1958) zurück. Das „stille“ bzw. „nicht ausgesprochene“ (tacit) Wissen liegt beispielsweise Handlungsroutinen, Erfahrungen, Einstellungen und Werten sowie Emotionen zugrunde (Polanyi, 1966). Nach Polanyi (1958, S. 87ff.) lassen sich das deklarative Wissen, das der Mehrheit bewusst ist, und das stille bzw. „unaussprechliche“ Wissen als entgegengesetzte Pole einer Dimension der Bewusstheit anordnen. Durch Prozesse der Wissenstransformation können weniger bewusste Wissensinhalte bewusstgemacht werden (Nonaka, 1994). Eine Methode der Wissenstransformation ist beispielsweise die Methode des Lauten Denkens (van Someren, Banard & Sandberg, 1994; Silby & Watts, 2015). Bewusstes und „stilles“ Wissen können aber auch in einer dynamischen Wechselseitigkeit zueinanderstehen und sich somit gegenseitig beeinflussen (Hildreth & Kimble, 2002).

Leighton (2017, S. 30) unterscheidet mit Bezug auf Ericsson und Simon (1993) drei Typen von Aufgabenstellungen für das Laute Denken. Aufgaben vom Typ 1 erfordern eine Verbalisation von etwas, das bereits versprachlicht ist, Aufgaben vom Typ 2 eine Verbalisation von Prozessen, die im Arbeitsgedächtnis ablaufen und leicht versprachlicht werden können, während Aufgaben vom Typ 3 Begründungen und Elaborationen erfordern, bei denen die Personen auf ihr Langzeitgedächtnis zurückgreifen. Erhebungsmethodisch schlägt Leighton (2017, S. 128) vor, die Erfassung simultan ablaufender Prozesse von der retrospektiven Erfassung höherer kognitiver Prozesse zu trennen (vgl. auch van Someren, Barnard & Sanberg, 1994). Ersteres könnte im Rahmen der Forschung zur Professionellen Wahrnehmung eher dem noticing (akzentuierend Typ 1- und Typ 2-Aufgabenstellung), letzteres eher dem reasoning (akzentuierend Typ 3-Aufgabenstellung) zugeordnet werden. Diese beiden Vorgehensweisen (simultanes Lautes Denken und retrospektives Interview) werden u. a. in der kognitiven Psychologie des Problemlösens eingesetzt (van Someren, Banard & Sandberg, 1994).

Wie bereits ausgeführt, können Wahrnehmung und Interpretation/Bewertung nicht streng unterschieden werden, da sie eher als ein Kontinuum gedacht werden müssen (vgl. auch Blömeke et al., 2015; König et al., 2014). Eine eindeutige Zuordnung der beiden Erhebungsmethoden (simultanes Lautes Denken, retrospektives Interview) zu den Prozessen der Wahrnehmung/Identifikation (noticing) der Interpretation/Begründung (reasoning) ist aber deshalb nicht möglich, weil auch beim simultanen Lauten Denken nicht nur Wahrnehmungs-, sondern auch Aspekte der Interpretation, d.h. der Bewertung geäußert werden; umgekehrt können bei einem retrospektiven Interview Wahrnehmungs- bzw. Identifikationsaspekte geäußert werden.

Die Videosequenzen, die bei der Erfassung der Professionellen Wahrnehmung Verwendung finden, können jedoch immer nur eine Auswahl aller möglichen Interaktionssituationen repräsentieren. Auch wenn die verwendeten Sequenzen von ExpertInnen und TeilnehmInnen als „authentisch“ (Seidel & Stürmer, 2014, S. 749) bzw. „typisch“ eingeschätzt werden (Seidel, Blomberg & Stürmer, 2010, S. 303), so stellt diese Authentizität bzw. Typikalität nicht sicher, ob eine Verwendung anderer Videosequenzen zu vergleichbaren Ergebnissen bei der Erfassung der Professionellen Wahrnehmung führt, die Messung somit „generalisierbar“ ist. Im Rahmen der Generalisierbarkeitstheorie (Cronbach, Gleser, Nanda, & Rajaratnam, 1972) stellen die Videosequenzen eine „Messfacette“ dar, deren Generalisierbarkeit ermittelt werden kann. Auch kann im Rahmen dieser Theorie bestimmt werden, wie hoch der Anteil anderer Messfacetten (z.B. theoretische Aspekte der Professionellen Wahrnehmung oder die Erhebungsmethode) an der Gesamtvarianz der Messwerte ausfällt.

2.2 Zusammenhänge zwischen Professioneller Wahrnehmung und biographischen Faktoren

Aus der Lehrerforschung ist bekannt, dass die Fähigkeit zur Professionellen Wahrnehmung mit der Berufserfahrung (bis hin zum „Expertenstatus“) ansteigt (Berliner, 2004). Mittlerweile stehen unterschiedliche Ausbildungswege und Ausbildungsebenen für die Tätigkeit in einer Kita zu Verfügung, die auch aufeinander aufbauend absolviert werden können (z.B. Fachschulausbildung, kindheitspädagogisches Bachelor- oder Masterstudium; Pasternack & Schulze, 2010). Unterschiedliche Ausbildungswege für die Ausbildung als pädagogische Fachkraft (z.B. Fachschule vs. Hochschule) sind außerdem mit unterschiedlichen Niveaus professionellen Wissens assoziiert (auf Sprache und Sprachförderung bezogenes Wissen; Strohmmer & Mischo, 2015). Daher ist auch anzunehmen, dass auch mit steigender Ausbildungsdauer und Ausbildungsniveau die Qualität der Professionellen Wahrnehmung zunimmt.

2.3 Konsequenzen für diese Studie

Da beide Methoden (simultanes Lautes Denken und retrospektives Interview) als geeignete Methoden zur Erfassung der Professionellen Wahrnehmung erscheinen, sollen diese beiden Methoden in dieser Studie eingesetzt werden. Außerdem soll untersucht werden, wie hoch der Anteil unterschiedlicher Messfacetten an der Erfassung der Professionellen Wahrnehmung ist. Dies soll auch Aufschluss darüber geben, in welchem Ausmaß die als Stimulusmaterial verwendeten Videosequenzen auf andere Sequenzen generalisierbar sind.

Als erste Hinweise auf die Validität der Erfassung sollen schließlich Zusammenhänge mit biographischen Faktoren (Berufserfahrung und Ausbildungsniveau) untersucht werden.

3. Fragestellungen

- Fragestellung 1: Kann die Professionelle Wahrnehmung intersubjektiv reliabel mit der Methode des Lauten Denkens (simultan) bzw. des retrospektiven Interviews erfasst werden?
- Fragestellung 2a: Wie hoch ist der Varianzanteil der unterschiedlichen Messfacetten an der gesamten Varianz?
- Fragestellung 2b: Ist die Messung generalisierbar über die konkreten videografierten dargebotenen Interaktionssituationen hinweg?
- Fragestellung 3: Gibt es einen positiven Zusammenhang zwischen der Berufserfahrung und der Ausbildungsdauer bzw. dem Ausbildungsniveau einerseits und der Professionellen Wahrnehmung andererseits?

4. Methode

4.1 Entwicklung des Stimulusmaterials: Aufzeichnung typischer Interaktionssequenzen und Pilotierung bei ExpertInnen

Folgende Anforderungen wurden an die Situationen formuliert, die später videografiert festgehalten und als Stimuli zur Erfassung der Professionellen Wahrnehmung eingesetzt werden: Die Situationen sollten sowohl Situationen einer Fachkraft mit einem einzelnen Kind als auch mit einer Gruppe enthalten, und sie sollten sowohl direkte wie indirekte Bildungssituationen darstellen. Außerdem sollten es Situationen sein, die typisch bzw. authentisch, wiederkehrend und möglichst allen späteren StudienteilnehmerInnen vertraut sind. Aufgrund der hohen Bedeutung sprachlicher Bildung in Kitas, die sowohl als Bildungsfeld in allen Bildungs- und Orientierungsplänen als auch in den Ausbildungsplänen

der Fachschulen und Hochschulen breit verankert sind (Mischo, 2016), fiel die Entscheidung, eine Einzel- und eine Gruppen-Vorlesesituation zu filmen (direkte Bildungssituationen). Außerdem sollte die Fachkraft-Kind-Interaktion in einer Essenssituation gefilmt werden, da diese Situation für den Kita-Alltag typisch ist und als indirekte Bildungssituation Bildungspotenzial enthält (Tietze & Viernickel, 2007; Gest et al., 2006). Schließlich sollte die Fachkraft-Kind-Interaktion in einer Freispielsituation gefilmt werden, da diese Situation ebenfalls als typische und indirekte Bildungssituation gelten kann (Wadepohl & Mackowiak, 2016). Die zu filmenden vier Sequenzen sollten unterschiedliche Fachkräfte darstellen. Deren Interaktionen mit Kindern sollten hinreichend fruchtbar sein, um Aspekte der drei CLASS-Domänen der Emotionalen Unterstützung, der Organisation des Kita-Alltags und der Lernunterstützung thematisieren zu können. Aus einer größeren Anzahl von Videosequenzen, die in Kitas im süd- und nordbadischen Raum gefilmt wurden, wurden in mehreren Durchgängen und Sichtungen durch die Projektgruppe schließlich insgesamt vier Sequenzen ausgewählt, die den oben genannten Kriterien entsprachen (Vorlesen einzeln, Vorlesen Gruppe, Mittagessen, Freispiel). In allen Fällen, in denen die Sequenz Kinder zeigte, wurde die Einwilligungserklärung der Eltern eingeholt. Das Alter der gefilmten Kinder lag zwischen 3 und 6 Jahren. Diese vier Videosequenzen mit einer Dauer zwischen 2 und 3,5 Minuten wurden drei unterschiedlichen Gruppen von ExpertInnen teilweise in Einzel-, teilweise in Gruppensitzungen vorgelegt ($N = 16$ insgesamt). Die ExpertInnen waren Personen, die viele Jahre im Bereich der Frühen Bildung als Lehrende und Forschende in entsprechenden Studiengängen, als FachberaterInnen oder Kita-Leitungen mit mehrjähriger Berufspraxis tätig waren. Die ExpertInnen wurden gebeten, während der Betrachtung der Sequenzen in Form des Lauten Denkens ihre spontanen Gedanken zu äußern (Noticing-Komponente). Retrospektiv wurden die ExpertenInnen explizit nach Bewertungen, Begründungen und Handlungsalternativen gefragt (Reasoning-Komponente). Die Äußerungen wurden auditiv aufgezeichnet. Diese ExpertInnen-Antworten dienten dazu, das Potenzial der Situation in Bezug auf die Evozierung Lauten Denkens abzuschätzen; außerdem wurden später einige konkrete ExpertInnen-Äußerungen als Ankerbeispiele für (höhere) Ausprägungen des zu entwickelnden Kategoriensystems übernommen. Dabei erwies es sich als vorteilhaft, wenn die ExpertInnen bei der ersten Präsentation die Videosequenzen stoppen konnten, um ihre spontanen Gedanken verbalisieren zu können und nicht die weitere Sequenz dadurch zu verpassen. Dieses Vorgehen wurde daher für die Datenerhebung übernommen. Die Pilotierung der Videosequenzen mit den ExpertInnen-Befragungen diente auch dazu, einen geeigneten Interviewleitfaden für das retrospektive Interview zu entwickeln.

Die vier Videosequenzen erwiesen sich als sehr gut geeignet, reichhaltige Inhalte in Form des Lauten Denkens zu evozieren. Anhand vorgegebener Rating-Items schätzten die ExpertInnen die vier Sequenzen (mit Abstrichen die Vorlesesituation in der Gruppe) als relativ typisch ein und gaben an, dass ihnen das Äußern von Kognitionen im spontanen Modus und das Beantworten der Fragen im retrospekti-

ven Modus (insbesondere bei den Vorlesesituationen) relativ leicht fiel (vgl. Tabelle 1).

Tabelle 1: Mittelwerte und Standardabweichungen der ExpertInnen-Ratings

| Videsequenz | M (SD) | | |
|-----------------|-------------|---|--|
| | Typikalität | Leichtigkeit der Beantwortung spontaner Modus | Leichtigkeit der Beantwortung retrospektiver Modus |
| Freispiel | 3.56 (0.53) | 3.40 (0.89) | 3.40 (0.89) |
| Mittagessen | 3.56 (0.53) | 3.33 (0.82) | 3.50 (1.00) |
| Vorlesen Einzel | 3.36 (0.92) | 4.00 (0.00) | 3.75 (0.50) |
| Vorlesen Gruppe | 3.00 (1.00) | 3.67 (0.58) | 3.67 (0.58) |

Anmerkungen. M = Mittelwert. SD = Standardabweichung. 1 = stimme nicht zu, 2 = stimme eher nicht zu, 3 = stimme eher zu, 4 = stimme zu.

4.2 Befragungsmethode zur Evokation der Professionellen Wahrnehmung

Um einerseits möglichst wenig reaktiv die simultanen Kognitionen zu erfassen (insbesondere bei der Noticing-Komponente), andererseits die interessierenden Aspekte (Reasoning-Komponente: Begründungen, angewandetes Wissen, Handlungsalternativen) mit Bezug zum Langzeitgedächtnis zu evozieren, wurden die vier Videosequenzen in zwei Durchgängen dargeboten. Das erste Mal wurden alle vier Sequenzen lediglich mit der Aufforderung zum Lauten Denken (und der Stopp-Möglichkeit der Videos) eingeleitet, das zweite Mal wurde jede Sequenz einzeln nochmals präsentiert und anschließend ein standardisiertes Interview geführt (retrospektives Interview v.a. zur Erfassung der Reasoning-Komponente). Dadurch wurde sichergestellt, dass sich die fokussierenden Fragen des Interviews in einer Videosequenz nicht auf die spontanen Kognitionen bei einer folgenden Videosequenz auswirkten. Beim retrospektiven Interview wurden die StudienteilnehmerInnen nochmals nach dem Geschehen in der Sequenz, nach Konkretisierungen, Begründungen, Bewertungen und Handlungsalternativen befragt. In weiteren Pilotierungsstudien mit vier angehenden und vier erfahrenen Kita-Fachkräften wurden das Prozedere und die Anwendbarkeit des Kategoriensystems erprobt. Dabei stellte sich bei der Nachbesprechung mit den TeilnehmerInnen heraus, dass diese beim retrospektiven Interview Inhalte deshalb nicht erneut äußerten, weil sie diese bereits beim Lauten Denken geäußert hatten. Um Redundanzen bei der Erhebung zu vermeiden, wurde für die Hauptstudie daher der Hinweis für die TeilnehmerInnen aufgenommen, dass sie beim retrospektiven Interview Inhalte nicht erneut äußern mussten.

4.3 Kategoriensystem zur Kodierung der Professionellen Wahrnehmung

Die Konstruktion des Kategoriensystems orientierte sich an gängigen Taxonomien zur Kategorisierung von LehrerInnenkognitionen (z.B. Peterson & Comeau, 1987; Tang & Watkins, 1994) bzw. Kognitionen von ErzieherInnen (Faas, 2013) und an speziell für die Professionelle Wahrnehmung eingesetzten Kategoriensystemen (Hamre et al., 2012; Schäfer & Seidel, 2015). In diesen Taxonomien werden unterschiedliche Aspekte der Professionellen Wahrnehmung unterschieden (z.B. Erinnern, Beschreiben, Interpretieren, theoriebasiertes Begründen, Peterson & Comeau, 1987; Tang & Watkins, 1994; van Es & Sherin, 2002, Kersting et al., 2012). Diese Aspekte unterscheiden sich in der Elaboration und dem Ausmaß theoretisch-konzeptueller Anreicherung. Inhaltlich können sich diese Aspekte der Professionellen Wahrnehmung auf unterschiedliche Merkmale der pädagogischen Interaktion beziehen. Zur Beschreibung dieser Merkmale der pädagogischen Interaktion kann auf die empirisch vielfach bewährten Domänen des CLASS Pre-K-Instruments (Emotionale Unterstützung, Organisation des Kita-Alltags, Lernunterstützung) zurückgegriffen werden. Diese inhaltlichen Domänen lassen sich als zusätzliche Dimension eines Kategoriensystems auffassen (Blömeke et al., 2015). Äußerungen bei der Erfassung der Professionellen Wahrnehmung, die sich auf die Aspekte Empathie, emotionale Zuwendung, Freundlichkeit, positives Klima, Ermunterung und Perspektivenübernahme der Fachkraft beziehen, wurden der Domäne *Emotionale Unterstützung* zugeordnet. Der Domäne *Organisation der Kita-Alltags* wurden Äußerungen zugeordnet, die sich auf die räumlich-organisatorische Struktur, den Beschäftigungs- und Beteiligungsgrad der Kinder und die Verhaltensregulation bezogen. Zur Domäne *Lernunterstützung* zählten Äußerungen, in denen kognitive Aktivierung, Feedback und die Förderung sprachlichen Lernens thematisiert wurden. Neben dieser theoretisch-deduktiven Vorgehensweise wurden aus den Pilotierungen mit ExpertInnen auch induktiv Kategorien gebildet (deduktiv-induktives Vorgehen, Mayring, 2010; Rustemeyer, 1992). Die Antworten wurden mit zunehmender Konkretisierung, Begründung und begrifflicher Abstraktion bis hin zu einer Integration von Erfahrungswissen und konzeptuellem Wissen im Sinne einer skalierenden Strukturierung kodiert (Mayring, 2010). Das Kategoriensystem für den Aspekt der Wahrnehmung zeigt Tabelle 2; die Kodierung erfolgte für jede der Domänen der Emotionalen Unterstützung, der Organisation des Kita-Alltags und der Lernunterstützung. Die Systeme für die Kodierung der Bewertung und der Handlungsalternativen waren strukturgleich aufgebaut.

Tabelle 2: Kategoriensystem für den Aspekt der Wahrnehmung

| Kategorienbezeichnung | Wert | |
|---|-------------------------------|---|
| Keine Nennung | 0 | |
| Nennung sequenzbezogener Oberflächenmerkmale | 1 | |
| Nennung einer relevanten Handlung und/oder Erfahrungswissen | Ohne begriffliche Abstraktion | 2 |
| | Mit begrifflicher Abstraktion | 3 |
| Konkretisierung einer relevanten Handlung und/oder Erfahrungswissen | Ohne begriffliche Abstraktion | 4 |
| | Mit begrifflicher Abstraktion | 5 |
| Konkretisierung einer relevanten Handlung und/oder Erfahrungswissen mit Begründung | Ohne begriffliche Abstraktion | 6 |
| | Mit begrifflicher Abstraktion | 7 |
| Konkretisierung und Begründung einer relevanten Handlung und/oder Erfahrungswissen durch konzeptuelles Wissen | 8 | |

Als Ankerbeispiele dienten Nennungen der Pilotierungsstudien.

Die *Messfacetten* des vorliegenden Designs für die weiteren Analysen bestanden somit außer den Personen (pädagogische Fachkräfte) aus den *vier Situationen der Videosequenzen* (Essenssituation, Vorlesen einzeln, Vorlesen Gruppe, Freispiel), den beiden *Erhebungsmethoden* bei jeder der Sequenzen (simultanes Lautes Denken vs. retrospektives Interview), den *theoretischen Aspekten der Professionellen Wahrnehmung* (Wahrnehmung, Bewertung, Handlungsalternativen) und den *drei CLASS-Domänen* innerhalb jeder dieser theoretischen Aspekte (Emotionale Unterstützung, Organisation des Kita-Alltags, Lernunterstützung). Die vier Situationen der Videosequenzen wurden als Zufallsfacette aufgefasst, die übrigen Facetten als feste Facetten, da sie auf theoretischen Überlegungen basieren und eine Generalisierung auf andere Ausprägungen dieser Facetten nicht beabsichtigt ist. In Tabelle 3 ist die Messfacette der vier Videosequenzen in den Zeilen, die Messfacetten der Erhebungsmethoden, der theoretischen Aspekte und der CLASS-Domänen in den Spalten dargestellt. Diese Messfacetten sind vollständig gekreuzt, d.h. bei allen Personen wurden alle möglichen Kombinationen aus Videosequenzen und Erhebungsmethoden realisiert und dabei die drei theoretischen Aspekte jeweils für alle drei CLASS-Domänen berücksichtigt. Jede Kombination entspricht in Tabelle 3 einer Zelle.

Tabelle 3: Messdesign

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---------------------------|-----|---|---|----------|---|---|--------------------------|---|---|-----|---|---|----------|---|---|---|---|---|
| Facette Sequenz | | Facette Erhebungsmethoden | | | | | | | | | | | | | | | | | | |
| Lautes Denken (simultan) | | | | | | | | | Retrospektives Interview | | | | | | | | | | | |
| Facette Aspekte der Professionellen Wahrnehmung | | | | | | | | | | | | | | | | | | | | |
| Wahrn | | | Bew | | | Alternat | | | Wahrn | | | Bew | | | Alternat | | | | | |
| Facette Domänen der Interaktion | | | | | | | | | | | | | | | | | | | | |
| E | O | L | E | O | L | E | O | L | E | O | L | E | O | L | E | O | L | E | O | L |
| VE | | | | | | | | | | | | | | | | | | | | |
| VG | | | | | | | | | | | | | | | | | | | | |
| Mitt | | | | | | | | | | | | | | | | | | | | |
| Frei | | | | | | | | | | | | | | | | | | | | |

Anmerkungen. VE = Vorlesesituation Einzeln. VG = Vorlesesituation Gruppe. Mitt = Mittagessenssituation. Frei = Freispielsituation Wahrn = Wahrnehmung, Bew = Bewertung, Alternat = Handlungsalternativen. E = Emotionale Unterstützung. O = Organisation des Kita-Alltags. L = Lernunterstützung.

Das gesamte Prozedere einschließlich der Anwendbarkeit des Kategoriensystems wurde schließlich bei vier NovizInnen und vier ExpertInnen (angehende vs. erfahrene Kita-Fachkräfte) pilotiert. Ziel war es dabei zu testen, ob die Kategorienformulierungen eindeutig nur einer der Kategorien zugeordnet werden kann (Exaktheit und Disjunktheit), ob alle spontanen Äußerungen einer Ausprägung des Kategoriensystems zugeordnet werden können, und ob das Kategoriensystem somit erschöpfend ist (Bortz & Döring, 2007, S. 255).

4.4 Durchführung

Da die Studie in ein umfangreicheres Forschungsprojekt (DFG-Projekt „Professionelle Wahrnehmung in KiTas ProWaK“) eingebettet war, bei der neben der Erfassung der Professionellen Wahrnehmung auch Fragebögen zu bearbeiten waren, und die TeilnehmerInnen außerdem selbst einen Vormittag im Gruppengeschehen der Kita mit dem CLASS Pre-K-Instrument beobachtet wurden, konnten die TeilnehmerInnen nur mit Hilfe einer finanziellen Aufwandsentschädigung von 100 € (anteilig für die Institution und die teilnehmende Person) gewonnen werden. Für die Teilnahme wurde deutschlandweit in sozialen Medien und Fachgruppen geworben. Ziel war es, eine möglichst große Varianz der Stichprobe hinsichtlich der Berufserfahrung und der Ausbildungswege (Fachschule, Hochschule) zu erzielen. Die Erhebung erfolgte durch geschulte Projektmitarbeiterinnen, die zu den Einzelerhebungen in die Kitas anreisten und die Videosequenzen computerunterstützt darboten. Die Äußerungen der Teilnehmerinnen wurden auditiv aufgezeichnet. Die Fragebögen wurden vor Ort bearbeitet.

Um Reihenfolgeeffekte zu vermeiden, wurden die vier Videosequenzen in der Hauptstudie in Form eines sequentiell ausbalancierten Lateinischen Quadrats verschiedenen StudienteilnehmerInnen in unterschiedlicher Reihenfolge dargeboten (Williams, 1949).

4.5 Stichprobe

An der Studie nahmen insgesamt 120 pädagogische Fachkräfte teil. Der Altersmittelwert lag bei 35.52 Jahren ($SD = 12.24$). Erwartungsgemäß war der überwiegende Anteil (107 Personen, 89.2 %) weiblichen Geschlechts, 13 Personen (10.8 %) waren männlichen Geschlechts. Alle Fachkräfte arbeiteten in Einrichtungen mit 3- bis 6-jährigen Kindern. Bei der Berufserfahrung lagen von 5 Personen fehlende Werte vor (4.2 %). Im Mittel waren die TeilnehmerInnen der Studie 10.57 Jahre berufstätig ($SD = 10.66$). Die Häufigkeiten des Ausbildungsniveaus sind in Tabelle 4 dargestellt.

Tabelle 4: Häufigkeiten der Ausbildungsstufen

| Ausbildungsstufe (ordinale Ausprägung) | Häufigkeit | Prozent |
|---|------------|---------|
| in Ausbildung an einer Fachschule (1) | 17 | 14.2 |
| im Studium an einer Hochschule (2) | 7 | 5.8 |
| Ausbildung abgeschlossen (3) | 69 | 57.5 |
| Studium unspezifisch abgeschlossen (4) | 10 | 8.3 |
| Studium spezifisch abgeschlossen (5) | 8 | 6.7 |
| Ausbildung abgeschlossen und im Studium (6) | 1 | 0.8 |
| Ausbildung und Studium unspezifisch abgeschlossen (7) | 4 | 3.3 |
| Ausbildung und Studium spezifisch abgeschlossen (8) | 4 | 3.3 |
| Gesamt | 120 | 100.0 |

Anmerkungen. Studium unspezifisch: ohne kindheitspädagogischen Bezug. Spezifisch: mit kindheitspädagogischem Bezug.

5. Ergebnisse

5.1 Fragstellung 1: Intersubjektivität der Kodierungen der Professionellen Wahrnehmung

Bevor die kodierten Kompetenzmaße für die Professionelle Wahrnehmung zur Beantwortung der Fragestellungen 2 und 3 verwendet werden können, muss geklärt werden, ob die Kodierungen intersubjektiv vergleichbar vorgenommen wurden.

Die Intersubjektivität der Kodierungen wurde durch eine Doppelkodierung von 10 Prozent des Datenmaterials (d.h. 12 von 120) durch zwei unabhängige

Kodiererinnen (eine „Hauptkodiererin“ und eine weitere Projektmitarbeiterin) sichergestellt. Beide Kodiererinnen hatten außerdem die Zertifizierung mit dem CLASS-Pre-K Beobachtungssystem erfolgreich absolviert, so dass die Bekanntheit mit der inhaltlichen Facette der CLASS-Domänen vorausgesetzt werden kann. Die Intraklassenkorrelationen der Kodierungen (two way mixed: Zufallsauswahl der Versuchspersonen und fixe Kodiererinnen) lagen für alle vier Sequenzen in einem guten Bereich (einzelne Maße: Vorlesesituation Einzeln ICC = .87, Vorlesesituation Gruppe ICC = .85, Mittagessenssituation ICC = .88, Freispielsituation ICC = .83). Zwei Drittel der Daten wurde von der Hauptkodiererin, ein Drittel von der weiteren Projektmitarbeiterin kodiert.

5.2 Fragestellung 2a: Varianzanteil der Messfacetten

Die simultane Abschätzung unterschiedlicher Varianzanteile bzw. Fehlerquellen von Messungen (z. B. Stimuli, Personen, Zeitpunkte, Items) leistet die Generalisierbarkeitstheorie (Cronbach et al., 1972; Webb & Shavelson, 2005). Je geringer der Messfehler der jeweiligen Fehlerquelle ausfällt, desto eher lässt sich die Messung über diese Fehlerquelle verallgemeinern. Dabei ist eine Äquivalenz-Annahme für Paralleltests, wie sie in der Klassischen Testtheorie postuliert wird, nicht nötig. Jeder gemessene Wert (z. B. die Qualität der Professionellen Wahrnehmung) stellt somit eine Stichprobe aus einem Universum möglicher Messungen unter verschiedenen Messbedingungen (Messfacetten) dar, wie beispielsweise die als Stimuli verwendeten Videosequenzen oder die Erhebungsmethode (simultanes Lautes Denken vs. retrospektives Interview). Der Erwartungswert der Messung über alle Messbedingungen wird in der Generalisierbarkeitstheorie als „universe score“ bezeichnet (Cronbach et al., 1972) und entspricht dem wahren Wert in der Klassischen Testtheorie. Über welche Messbedingungen bzw. Facetten generalisiert werden soll, und welche Facetten als theoretisch begründet und daher als „fest“ angenommen werden, ist auf Grund theoretischer Überlegungen zu entscheiden. Empirisch kann dann mit Hilfe der Generalisierungstheorie zunächst bestimmt werden, wie groß der Varianz- bzw. Fehleranteil der einzelnen Messfacetten ausfällt (Fragestellung 2a). Im Fokus der Generalisierbarkeitsstudie steht die Facette der ausgewählten Videosequenzen, da diese zwar nur eine Stichprobe aus einem theoretisch unendlichen Universum möglicher Interaktionssituationen darstellen, sie aber zur Messung des „universe score“ herangezogen werden (vgl. unten Fragestellung 2b). Da es für die anderen Messfacetten nur schwer möglich erscheint, ein theoretisch unendliches Universum möglicher Ausprägungen zu konstruieren, wird für diese keine Generalisierbarkeit angestrebt, sondern sie werden theoretisch begründet. Die Generalisierbarkeit der Messung ist eine notwendige, aber keine hinreichende Bedingung für die Validität.

Da die Messfacetten als zufällige (Personen, Videosequenzen) bzw. fixe Effekte (die weiteren Facetten) fungieren, entspricht die Auswertungslogik einer varianzanalytischen Zerlegung der Gesamtvarianz in die Varianzen der einzel-

nen Messfacetten mit fixen und zufälligen Faktoren. Zur Auswertung wurde das Programm EduG in der Version 6.1.e verwendet (Swiss Society for Research in Education Working Group, 2010; Cardinet, Johnson & Pini, 2010). Um einen ersten Überblick über die Daten zu erhalten, wurden zunächst die deskriptiven Maße für die Kompetenz-Scores für die Professionelle Wahrnehmung für alle Messfacetten berechnet (siehe Tabelle 5).

Tabelle 5: Mittelwerte für die Messfacetten

| Messfacette der Professionellen Wahrnehmung | | <i>M</i> (<i>SD</i>) | Min | Max |
|--|-------------------------------|--------------------------|-----|-----|
| Videsequenz | Freispiel | 0.94 (1.47) | 0 | 7 |
| | Mittagessen | 0.99 (1.46) | 0 | 6 |
| | Vorlesen Einzel | 0.73 (1.31) | 0 | 6 |
| | Vorlesen Gruppe | 0.83 (1.41) | 0 | 6 |
| Modus der Erfassung | spontan | 0.90 (1.36) | 0 | 6 |
| | retrospektiv | 0.85 (1.46) | 0 | 7 |
| Aspekt der Professionellen Wahrnehmung | Wahrnehmung | 1.17 (1.51) | 0 | 6 |
| | Bewertung | 0.91 (1.46) | 0 | 7 |
| | Handlungsalternativen | 0.54 (1.19) | 0 | 7 |
| Handlungsdomänen | Emotionale Unterstützung | 0.73 (1.33) | 0 | 7 |
| | Organisation des Kita-Alltags | 1.02 (1.47) | 0 | 6 |
| | Lernunterstützung | 0.87 (1.43) | 0 | 7 |
| Gesamtmittelwert (grand mean der Messfacetten) | | 0.87 (0.08) ^a | | |

Anmerkungen: *M* = Mittelwert. *SD* = Standardabweichung. Min = Minimum. Max = Maximum.

^aStandardfehler des grand mean.

Die Mittelwerte sind die für jede Stufe einer Messfacette über alle anderen Messfacetten gemittelten Werte und entsprechen damit in einem varianzanalytischen Design den Mittelwerten für die Stufen der unabhängigen Variablen (Haupteffekte; vgl. Tabelle 3). Zwar liegt der höchste theoretische Wert des Kategoriensystems bei 8, aufgrund der vielen Nicht-Nennungen bei einzelnen Facetten (Wert 0) ergeben sich jedoch insgesamt (d. h. für jede Person gemittelt über alle Zellen des Designs in Tabelle 3) vergleichsweise geringe Ausprägungen.

Hinsichtlich der vier Sequenzen ergaben sich etwas höhere Scores in der Professionellen Wahrnehmung bei den dargebotenen Freispiel- und Mittagessen-Videos als bei den Vorlesesituationen. Außerdem wurden beim spontanen Modus des Lauten Denkens etwas höhere Scores bei der Wahrnehmung erreicht als beim retrospektiven Interview-Modus. Unterschiede zeigten sich auch in den Scores in Bezug auf die theoretischen Aspekte der Professionellen Wahrnehmung: Bei der Identifikation traten die höchsten kodierten Werte auf, geringere Werte bei der Bewertung und mit Abstand die geringsten Werte bei den geäußerten Handlungsalternativen.

In Bezug auf die Interaktionsdomänen der CLASS zeigten die Personen die höchsten Scores bei der Organisation des Kita-Alltags, gefolgt von den Scores bei der Lernunterstützung.

Die Varianzanteile der Messfacetten sind in Tabelle 6 dargestellt.

Tabelle 6: Varianzanteile der Messfacetten

| Quelle der Varianz | Varianz ^a | % | SE |
|--------------------|----------------------|-------|------|
| P | 0.26 | 6.2 | 0.05 |
| S | 0.01 | 0.3 | 0.01 |
| M | -0.01 | 0.0 | 0.01 |
| A | 0.05 | 1.3 | 0.07 |
| D | 0.01 | 0.1 | 0.02 |
| PS | 0.08 | 1.8 | 0.01 |
| PM | 0.49 | 11.5 | 0.07 |
| PA | 0.12 | 2.9 | 0.02 |
| PD | 0.03 | 0.7 | 0.01 |
| SM | 0.01 | 0.1 | 0.01 |
| SA | 0.04 | 1.0 | 0.02 |
| SD | 0.05 | 1.2 | 0.03 |
| MA | 0.01 | 0.1 | 0.01 |
| MD | 0.01 | 0.0 | 0.01 |
| AD | 0.01 | 0.1 | 0.01 |
| PSM | 0.16 | 3.8 | 0.01 |
| PSA | 0.27 | 6.4 | 0.02 |
| PSD | 0.17 | 4.1 | 0.01 |
| PMA | 0.14 | 3.3 | 0.03 |
| PMD | 0.03 | 0.6 | 0.01 |
| PAD | 0.01 | 0.4 | 0.01 |
| SMA | 0.02 | 0.6 | 0.01 |
| SMD | 0.01 | 0.2 | 0.01 |
| SAD | 0.01 | 0.4 | 0.01 |
| MAD | -0.01 | 0.0 | 0.01 |
| PSMA | 0.46 | 10.8 | 0.03 |
| PSMD | 0.32 | 7.5 | 0.02 |
| PSAD | 0.50 | 11.8 | 0.02 |
| PMAD | 0.04 | 0.9 | 0.02 |
| SMAD | 0.01 | 0.2 | 0.01 |
| PSMAD | 0.93 | 21.7 | 0.03 |
| Total | | 100.0 | |

Anmerkungen. P = Person, S = Sequenz (1 bis 4), M = Modus (spontan vs. retrospektiv), A = Aspekt der prof. Wahrnehmung (Wahrnehmung, Bewertung, Handlungsalternativen), D = Domäne (Emotionale Unterstützung, Organisation des Kita-Alltags, Lernunterstützung), SE = Standardfehler der geschätzten (korrigierten) Varianzkomponente.

^aKorrigiert nach Whimbey, Vaughan und Tatsuoka (1967).

Von den Hauptmessfacetten entfällt der größte Varianzanteil auf die fünffache Wechselwirkung der Facetten (21.7 %), gefolgt von der dreifachen Wechselwirkung Personen x Szenario x theoretischer Aspekt der Professionellen Wahrnehmung (11.8 %) und der zweifachen Wechselwirkung Personen x Erhebungsmethode (11.5 %).

5.3 Fragestellung 2b: Generalisierbarkeit der Messung – Generalisierbarkeitskoeffizient und Facettenanalyse

Im Rahmen der Generalisierbarkeitstheorie gibt der Generalisierbarkeitskoeffizient G den Anteil der wahren Varianz an der Gesamtvarianz der Messwerte an. Dabei kann ein relativer von einem absoluten G -Koeffizienten unterschieden werden. Der relative G -Koeffizient bezieht sich auf eine Messung auf einer relativen Skala, bei der Unterschiede zwischen den Personen im Vordergrund stehen. Der absolute G -Koeffizient bezieht sich auf eine Skala, bei der absolute Messungen thematisch sind, wie beispielsweise das Erreichen eines absoluten Wertes bei einem lernzielorientierten Test. Da in weiteren Analysen Korrelationsanalysen durchgeführt werden, ist der relative G -Koeffizient von Bedeutung. Für den Gesamtscore beträgt der relative G -Koeffizient $G = .93$. Werte über $.80$ und können als zuverlässig gelten (Swiss Society for Research in Education Working Group, 2010, S. 23).

Außer der Generalisierbarkeit des Gesamtscores interessieren die G -Koeffizienten der einzelnen Messfacetten (Facettenanalyse). Die Koeffizienten der Facettenanalyse geben an, wie sich die G -Koeffizienten im Vergleich zum G -Koeffizienten des Gesamtscores ($G = .93$) verändern würden, wenn diese Ausprägung der Messfacette *eliminiert würde* (Swiss Society for Research in Education Working Group, 2010, S. 28). Dies ist v. a. bei fixen Facetten (hier also nicht für die zufälligen Facetten der Videosequenzen) relevant (Swiss Society for Research in Education Working Group, 2010, S. 28).

Tabelle 7: Ergebnisse der Facettenanalyse

| Messfacette der Professionellen Wahrnehmung | | Relativer G-Koeffizient bei <i>Elimination</i> der Stufe der Messfacette (G-gesamt = .93) |
|---|-------------------------------|---|
| Videsequenz | Freispiel | .89 |
| | Mittagessen | .90 |
| | Vorlesen Einzel | .92 |
| | Vorlesen Gruppe | .92 |
| Modus der Erfassung | spontan | .95 |
| | retrospektiv | .81 |
| Aspekt der Professionellen Wahrnehmung | Wahrnehmung | .80 |
| | Bewertung | .77 |
| | Handlungsalternativen | .81 |
| Handlungsdomänen | Emotionale Unterstützung | .88 |
| | Organisation des Kita-Alltags | .86 |
| | Lernunterstützung | .88 |

Offenbar würde ein Verzicht auf den spontanen Antwortmodus zu einer leichten Steigerung des G-Koeffizienten beitragen. Ein Design ohne den Aspekt der Bewertung würde dagegen den G-Koeffizienten deutlich niedriger ausfallen lassen. Da der G-Koeffizient aber insgesamt sehr hoch ausfällt, besteht letztlich keine Veranlassung, auf die hier umgesetzten (fixen) Messfacetten in zukünftigen Studien zu verzichten.

5.4 Fragestellung 3: Zusammenhang mit Ausbildungsstufe und Berufserfahrung

Da die fehlenden Werte bei der Variablen Berufserfahrung bei fünf Personen vollständig zufällig verteilt sind (Little's MCAR Test: Chi-Quadrat = 0.48, $df = 71$, $p = 1$), wurden die Missings zur Berechnung der Korrelationen mit dem Expectation Maximization Algorithmus in SPSS geschätzt (EM-Algorithmus).

Tabelle 8: Korrelationen der Messfacetten mit Ausbildungsstufe und Berufserfahrung

| Messfacette der Professionellen Wahrnehmung | | Korrelationen ^a (N = 120) | |
|---|-------------------------------|---|-----------------|
| | | Ausbildungsstufe | Berufserfahrung |
| Gesamtscore | | .22** | .19* |
| Videsequenz | Freispiel | .13+ | .22** |
| | Mittagessen | .19* | .21* |
| | Vorlesen Einzeln | .20* | .12 |
| | Vorlesen Gruppe | .14+ | .14 |
| Modus der Erfassung | spontan | .14+ | .13 |
| | retrospektiv | .21* | .17* |
| Aspekt der Professionellen Wahrnehmung | Wahrnehmung | .17* | .11 |
| | Bewertung | .10 | .23** |
| | Handlungsalternativen | .18* | .14 |
| Handlungsdomänen | Emotionale Unterstützung | .14+ | .20* |
| | Organisation des Kita-Alltags | .22** | .19* |
| | Lernunterstützung | .16* | .15 |

Anmerkungen. Berufserfahrung mit EM-Schätzung der Missings.

^aKorrelationen mit Ausbildungsstufe: Spearmans rho. Korrelationen mit Berufserfahrung: Spearmans Produkt-Moment-Korrelationen.

+ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. Jeweils einseitige Testung aufgrund gerichteter Fragestellung 3.

Es ergeben sich statistisch bedeutsame Zusammenhänge zwischen dem Gesamtwert der Professionellen Wahrnehmung und der Ausbildungsstufe sowie der Berufserfahrung. Die Berufserfahrung scheint dabei eher mit der Professionellen Wahrnehmung in den Sequenzen *Freispiel* und *Mittagessen* zusammenzuhängen, während die Ausbildungsstufe auch mit der Professionellen Wahrnehmung in den *Vorlesesituationen* zusammenhängt. Sowohl Ausbildungsstufe als auch Berufserfahrung zeigen Zusammenhänge mit der retrospektiven Erfassung der Professionellen Wahrnehmung. Bei der Messfacette des *Aspekts der Professionellen Wahrnehmung* (Wahrnehmung/Identifikation, Interpretation bzw. Bewertung, Handlungsalternativen) ergeben sich unterschiedliche Zusammenhangsmuster mit der Ausbildungsstufe vs. der Berufserfahrung. Interessanterweise korreliert nur die Ausbildungsstufe (nicht jedoch die Berufserfahrung) mit der Professionellen Wahrnehmung der Lernunterstützung.

6. Diskussion

Insgesamt scheint es mit dem vorgestellten Verfahren gelungen zu sein, ein intersubjektiv anwendbares Instrument zu konstruieren, das die Professionelle Wahrnehmung in relativ typischen Interaktionssituationen zu erfassen erlaubt (Fragestellung 1). Das Ergebnis, dass die Scores der Professionellen Wahrnehmung in den Vorlesesituationen etwas geringer ausfallen als in den Freispiel- und Mittagessenssituationen mag etwas überraschen, da Vorlesesituationen eigentlich als typische und primäre Bildungssituationen gelten können, für die theorieangereichertes Wissen bei der Wahrnehmung entsprechender Episoden zur Verfügung stehen sollte. Es ist jedoch zu berücksichtigen, dass gerade in der Einzelsituation einige Aspekte (beispielsweise der Organisation des Kita-Alltags) nicht so salient sind und daher auch seltener genannt wurden. Auch der Befund, dass im spontanen Antwortmodus des Lauten Denkens etwas höhere Scores erzielt wurden als bei dem retrospektiven Interview, ist erklärungsbedürftig, da die Vermutung angestellt werden könnte, dass eher bei einer fokussierteren Befragung konzeptuelle Anreicherungen genannt werden können. Zumindest teilweise könnte dieses Ergebnis jedoch darauf zurückzuführen sein, dass die TeilnehmerInnen beim retrospektiven Interview darauf hingewiesen wurden, dass Nennungen, die sie bereits beim simultanen Lauten Denken angeführt haben, nicht nochmals äußern müssen. Um die Effekte unterschiedlicher Erhebungsmethoden genauer zu untersuchen, wäre daher ein Between-Design eher angemessen. Die etwas höheren Werte beim simultanen Lauten Denken könnten außerdem auch dadurch verursacht sein, dass bereits bei dieser Methode Bewertungs- und Begründungsaspekte genannt wurden, da bei der Identifikation relevanter Interaktionsmerkmale Tiefen- und keine Oberflächenmerkmale relevant sind, die bereits Interpretationsaspekte enthalten (z. B. die Äußerung: [die Fachkraft „(...) regt zum Nachdenken an“, oder „(...) stellt eine Bindung zu den Kindern her“). Dies könnte darauf hindeuten, dass die in der Forschung zur Professionellen Wahrnehmung unterschiedenen Komponenten der Identifikation (noticing) und der Bewertung bzw. Begründung (reasoning) theoretisch fließende Übergänge aufweisen und – zumindest mit den hier gewählten Erhebungsmethoden – erhebungsmethodisch nur schwer zu trennen sind. Dies ist mit der theoretischen Annahme vereinbar, dass sich Wahrnehmung (noticing) und Interpretation (reasoning) eher als ein Kontinuum darstellen lassen (Blömeke et al., 2015). Möglicherweise bieten andere Erhebungsverfahren, wie beispielsweise das Eye-Tracking, weitere Möglichkeiten der Verortung von Wahrnehmungsaspekten auf dieser theoretisch postulierten Dimension, wobei sich dann die Frage stellt, ob die ProbandInnen das in Augenschein Genommene wirklich wahrnehmen und die Tiefenstruktur der sozialen Interaktion erkennen. Möglicherweise wäre erst eine Kombination dieser Methoden in der Lage, die jeweiligen Nachteile bestimmter Methoden zu kompensieren.

Aufgrund des relativ hohen Aufwands der Kodierung offener Äußerungen stellt sich außerdem die Frage, ob vergleichbare Ergebnisse nicht auch bei einer ökonomischeren Methode erzielt werden könnten.

mischeren Methode (z.B. Rating-Items) erzielt werden können. Trotz der theoretischen Einwände gegen die Vorgabe von Items ist dies letztlich eine empirisch zu klärende Frage.

Dass bei der Professionellen Wahrnehmung die Scores in der Domäne der Lernunterstützung höher ausfielen als in der Domäne der Emotionalen Unterstützung, steht im Gegensatz zu Beobachtungsstudien, bei denen übereinstimmend die niedrigsten Werte im Verhalten der CLASS-Domäne der Lernunterstützung auftraten (Bihler et al., 2018; Stuck, Kammermeyer & Roux, 2016; von Suchodoletz, Fäsche, Gunzenhauer & Hamre, 2014; Wildgruber, Wertfein, Wirts, Kammermeier & Danay, 2016; Wirts et al., 2018). Allerdings war die Lernunterstützung in den ausgewählten Videos vermutlich auch relativ salient, da diese Sequenzen auch daraufhin ausgewählt wurden, ob sie nach Ansicht der ExpertInnen lernunterstützende Momente enthielten.

Da jedoch keine Hypothesen über die Unterschiede zwischen den Messfacetten formuliert wurden, wurde auf eine inferenzstatistische Prüfung der Unterschiedlichkeit der Scores verzichtet. Die berichteten Unterschiede sind daher explorativ zu interpretieren.

Auch sind die Varianzanteile für die Haupteffekte der Messfacetten (Szenarien, Erhebungsmodus, theoretischer Aspekt der Professionellen Wahrnehmung) im Gegensatz zum Varianzanteil mehrfacher Wechselwirkungen relativ gering (Fragestellung 2a). Eine theoretisch-inhaltliche Interpretation des Effekts dieser Wechselwirkungen fällt schwer, jedoch ist die Generalisierbarkeit – zumindest der Zufallsfacette der Videosequenzen – relativ hoch (Fragestellung 2b). Da die anderen Messfacetten feste (und nicht zufällige) Faktoren darstellten, war eine Generalisierung über die hier realisierten Ausprägungen hinaus nicht Gegenstand der Studie. Es muss daher zukünftigen Studien vorbehalten bleiben, ob andere Erhebungsmodi, andere Aspekte der Professionellen Wahrnehmung oder andere Interaktionsdomänen – gerade auch in Wechselwirkung zueinander – zu anderen Ergebnissen bei der Erfassung der Professionellen Wahrnehmung führen.

Als ein erster Hinweis für die Validität der Messung können die Zusammenhänge mit der Ausbildungsstufe und der Berufserfahrung gelten (Fragestellung 3). Das Ergebnis, dass die Ausbildungsstufe mit der Professionellen Wahrnehmung – im Gegensatz zur Berufserfahrung – auch in den Vorlesesituationen korreliert, könnte darauf zurückzuführen sein, dass im Zuge der stark gestiegenen Bedeutung der Sprachförderung dieser Aspekt ein immer wichtigerer Bestandteil der Ausbildungs- und Orientierungspläne geworden ist. Dass die Ausbildungsstufe (nicht jedoch die Berufserfahrung) mit den gemittelten Scores der Professionellen Wahrnehmung der Lernunterstützung kovariiert, könnte damit zusammenhängen, dass Themen der Lernunterstützung (z.B. Sprachförderung) mit zunehmender Ausbildungsstufe intensiver behandelt werden. Vorliegende Befunde über Zusammenhänge zwischen Ausbildungsstufe und sprachbezogenem Wissen stützen diese Annahme (Strohmer & Mischo, 2015). Dieses Wissen kann bei der Professionellen Wahrnehmung (gerade bei den zwei Videosequenzen zur Sprachförderung) herangezogen werden. Der Befund, dass die Professionelle

Wahrnehmung in der Domäne der Emotionalen Unterstützung mit der Berufserfahrung korreliert, könnte so interpretiert werden, dass – ausbildungsunabhängig – die Relevanz emotionaler Interaktionsaspekte für die Fachkräfte wichtiger wird. Diese Interpretationen müssen aber letztlich spekulativ bleiben. Die Befunde stehen jedoch insgesamt in Einklang mit zentralen Ergebnissen der Expertiseforschung, die der Berufserfahrung und dem Wissen eine wichtige Funktion für die Professionelle Wahrnehmung zumessen (Berliner, 2004). Die Entwicklung von Expertise trägt somit nicht zur Prozeduralisierung (Blömeke et al., 2010) von Handlungsroutinen, sondern auch von Wahrnehmungsroutinen bei (Landy, 2018, S. 152), die eng mit den Handlungsroutinen verschränkt sein können (Gobet, 2005).

Es gehört daher zu den Einschränkungen dieser Studie, dass der Fokus auf die Reliabilität und Generalisierbarkeit und nur in Ansätzen auf die Validität der Messungen gelegt werden konnte. Es stellt sich jedoch auch die Frage, welche Validitätskriterien für welche Validitätsaspekte sinnvollerweise anzusetzen sind. Auch bei einer konvergenten Validierung mit anderen, weniger reaktiven Verfahren dürfte es schwierig zu bestimmen sein, welches Verfahren als „valideres“ Verfahren anzusetzen ist.

In der Lehramtsausbildung wird die Schulung der Professionellen Wahrnehmung bereits erfolgreich eingesetzt (Sherin & van Es, 2005; Blomberg, Renkl, Sherin, Borko & Seidel, 2013). Zusammenhänge mit dem beobachtbaren Verhalten werden dabei jedoch selten untersucht. Da im Gesamtprojekt auch das Verhalten der TeilnehmerInnen (mit der CLASS Pre-K) erfasst wurde, müssen weitere Analysen zeigen, in welchen CLASS-Domänen des Verhaltens sich Zusammenhänge mit den Messfacetten der Professionellen Wahrnehmung empirisch nachweisen lassen. Sollten sich hier Kovariationen zeigen, wäre dies ein weiterer Hinweis dafür, dass die Professionelle Wahrnehmung eine wichtige Bedingung für erfolgreiches professionelles Handeln in einer Kita darstellt (Pianta et al., 2014). Auch müssen erst die Korrelationsanalysen mit dem Verhalten zeigen, ob die relativ aufwändige Erfassung in einem offenen Antwortmodus einer ökonomischeren Erfassung mit Rating-Items (und der Berechnung von Abweichungen zu ExpertInnen-Einschätzungen, z.B. Schäfer & Seidel, 2015) überlegen ist. Die Verhaltensrelevanz vorausgesetzt, erlaubt es die Ausdifferenzierung nach den hier unterschiedlichen Messfacetten, entsprechende Trainingsmaßnahmen der Professionellen Wahrnehmung spezifischer zu konzipieren und zu evaluieren.

Danksagung

Wir danken der Deutschen Forschungsgemeinschaft für die finanzielle Förderung dieses Forschungsprojekts (GZ: MI 690/4-1).

Literatur

- Anders, Y. (2012). *Modelle professioneller Kompetenzen für frühpädagogische Fachkräfte. Aktueller Stand und ihr Bezug zur Professionalisierung*. München: Aktionsrat Bildung.
- Ahnert, L. & Gappa, M. (2013). Bindung und Beziehungsgestaltung in öffentlicher Kleinkindbetreuung – Auswirkungen auf die Frühe Bildung. In H. R. Leu & A. von Behr (Hrsg.), *Forschung und Praxis der Frühpädagogik. Profiwissen für die Arbeit mit Kindern von 0-3 Jahren* (S. 110–121). München: Ernst Reinhardt.
- Bannert, M. (2004). Erfassung von metakognitiv-strategien Lernaktivitäten. In M. Wosintza, A. Frey, & R. Jäger (Hrsg.), *Lernprozess, Lernumgebung und Lern-diagnostik*. (S.375-391). Landau: Verlag Empirische Pädagogik.
- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24(3), 200–213.
- Berliner, D.C. (1987). In pursuit of the expert pedagogue. *Educational Researcher*, 15, 5–13.
- Bihler, L. M., Agache, A., Kohl, K., Willard, J. A. & Leyendecker, B. (2018). Factor analysis of the classroom assessment scoring system: Replicates the three domain structure and reveals no support for the bifactor model in German preschools. *Frontiers in Psychology*, 9, 11–13.
- Blomberg, G., Renkl, A., Sherin, M. G., Borko, H. & Seidel, T. (2013). Five research-based heuristics for using video in preservice teacher education. *Journal for Educational Research Online*, 5(1), 90–114.
- Blömeke, S., Gustafsson, J.-E. & Shavelson, R. J. (2015). Beyond dichotomies. Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223, 3–13.
- Blömeke, S., Kaiser, G. & Lehmann, R. (2010). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Bortz, J. & Döring, N. (2007). *Forschungsmethoden und Evaluation*. Heidelberg: Springer.
- Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Bern: Huber.
- Burchinal, M., Vandergrift, N., Pianta, R. & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in prekindergarten programs. *Early Childhood Research Quarterly*, 25(2), 166–176.
- Cardinet, J., Johnson, S. & Pini, G. (2010). *Applying generalizability theory using EduG*. New York, NY: Routledge.
- Clark, C. M. & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching: A project of the American Educational Research Association* (3. Aufl., S. 255–296). New York, NY: Macmillan Library Reference.
- Cronbach, L. J., Gleser, G. C., Nanda, H. & Rajaratnam, N. (1972). *The dependability of behavioral measurements*. New York: Wiley.
- Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Hrsg.), *The Cambridge handbook of expertise and expert performance* (S. 223–241). New York, NY: Cambridge University.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT.
- Faas, S. (2013). *Berufliche Anforderungen und berufsbezogenes Wissen von Erzieherinnen. Theoretische und empirische Rekonstruktionen*. Wiesbaden: VS.

- Fröhlich-Gildhoff, K., Nentwig-Gesemann, I. & Pietsch, S. (2011). *Kompetenzorientierung in der Qualifizierung frühpädagogischer Fachkräfte*. München: Weiterbildungsinstitut für Frühpädagogische Fachkräfte (WiFF).
- Fröhlich-Gildhoff, K., Nentwig-Gesemann, I., Pietsch, S., Köhler, L. & Koch, M. (Hrsg.). (2014). *Kompetenzentwicklung und Kompetenzerfassung in der Frühpädagogik: Konzepte und Methoden incl. Begleit-CD*. Freiburg: FEL-Verlag Forschung-Entwicklung-Lehre.
- Gest, S. D., Holland-Coviello, R., Welsh, J. A., Eicher-Catt, D. L. & Gill, S. (2006). Language development subcontexts in HeadStart classrooms: Distinctive patterns of teacher talk during free play, mealtime and book reading. *Early Education and Development, 17*, 293–315.
- Gobet, F. (2005). Chunking models of expertise: Implications for education. *Applied Cognitive Psychology, 19*, 183–204.
- Gold, B., Hellermann, C. & Holodynski, M. (2016). Professionelle Wahrnehmung von Klassenführung – Vergleich von zwei videobasierten Erfassungsmethoden. In K. Schwippert & D. Prinz (Hrsg.), *Der Forschung – Der Lehre – Der Bildung: Aktuelle Entwicklungen der Empirischen Bildungsforschung* (S. 103–118). Münster: Waxmann
- Hamre, B. & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*(5), 949–967.
- Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., Howes, C. LaParo, K. & Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge and observed practice. *American Educational Research Journal, 49*, 88–123.
- Hildreth, P. M. & Kimble, C. (2002). The duality of knowledge. *Information Research, 8*(1), 27–39.
- Hogan, T. & Rabinowitz, M. (2009). Teacher expertise and the development of a problem representation. *Educational Psychology, 29*, 153–169.
- Jamil, F. M., Sabol, T. J., Hamre, B. K. & Pianta, R. C. (2015). Assessing teachers' skills in detecting and identifying effective interactions in the classroom: Theory and measurement. *The Elementary School Journal, 115*(3), 407–432.
- Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching. *Mathematics Educational and Psychological Measurement, 68*(5), 845–861.
- Kersting, N., Givvin, K., Thompson, B., Santaga, R. & Stigler, J. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal, 49*(3), 568–589.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A. & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education, 38*, 76–88.
- Krull, E., Oras, K. & Sisask, S. (2007). Differences in teachers' comments on classroom events as indicators of their professional development. *Teaching and Teacher Education, 23*, 1038–1050.
- Lamb, M. E. (1998). Nonparental child care: Context, quality, correlates, and consequences. In Sigel, I. E. (Hrsg.), *Child psychology in practice, Volume 4* (S. 73–133). New York, NY: Wiley.
- Landy, D. (2018). Perception in expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt & A. M. Williams (Hrsg.), *The Cambridge handbook of expertise and expert performance* (S. 151–164). Cambridge University.
- Leighton, J. P. (2017). *Using think-aloud interviews and cognitive labs in educational research*. Oxford: Oxford University.

- Mackowiak, K., Wadepohl, H. & Bossart, S. (2014). *Analyse von Kompetenzen von pädagogischen Fachkräften im Freispiel und in Bildungsangeboten*. In D. Kucharz, K. Mackowiak, S. Zirolì, A. Kauertz, E. Rathgeb-Schnierer & M. Dieck (Hrsg.), *Professionelles Handeln im Elementarbereich (PRIMEL): Eine deutsch-schweizerische Videostudie* (S. 179–204). Münster: Waxmann.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz.
- Mischo, C. (2016). Subjektiver Kompetenzgewinn und Wissenszuwachs bei frùhpädagogischen Fachkräften unterschiedlicher Ausbildungsprofile. *Zeitschrift für Erziehungswissenschaft*, 19(3), 577–597.
- NICHHD. (2002). Child-care structure->process->outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, 13(3), 199–206.
- Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organisation Science*, 5(1), 14–37.
- Pasternack, P. & Schulze, H. (2010). *Die frùhpädagogische Ausbildungslandschaft. Strukturen, Qualifikationsrahmen und Curricula. Gutachten für die Robert-Bosch-Stiftung (HoF-Arbeitsbericht 2'2010)*. Wittenberg: Institut für Hochschulforschung (HoF) an der Martin-Luther-Universität Halle-Wittenberg.
- Peterson, P. L. & Comeaux, M. A. (1987). Teachers' schemata for classroom events: The mental scaffolding of teachers' thinking during classroom instruction. *Teaching and Teacher Education*, 3, 319–331.
- Pianta, R. C., Burchinal, M., Jamil, F. M., Sabol, T., Grimm, K., Hamre, B. K., Downer, J., LoCasala-Crouch, J. & Howes, C. (2014). A cross-lag analysis of longitudinal associations between preschool teachers' instructional support identification skills and observed behavior. *Early Childhood Research Quarterly*, 29(2), 144–154.
- Pianta, R. C., La Paro, K. M. & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS): Manual PRE-K-K*. Baltimore: Paul H. Brookes.
- Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. Chicago, IL: University of Chicago.
- Polanyi, M. (1966). *The Tacit Dimension*. Chicago, IL: University of Chicago.
- Rustemeyer, R. (1992). *Praktisch-methodische Schritte der Inhaltsanalyse*. Münster: Aschendorff.
- Schäfer, S. & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. *Journal of Educational Research Online*, 7, 34–58.
- Schweer, M., Thies, B. & Lachner, R. B. (2016). Soziale Wahrnehmung und unterrichtliches Handeln. In M. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (S. 121–146). Wiesbaden: Springer.
- Seidel, T., Blomberg, G. & Stürmer, K. (2010). „Observer“ – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. *Zeitschrift für Pädagogik (Beiheft)*, 56, 296–306.
- Seidel, T. & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*, 51(4), 739–771.
- Sherin, M. G. & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education*, 13, 475–491.
- Silby, A. & Watts, M. (2015). Making the tacit explicit: children's strategies for classroom writing. *British Educational Research Journal*, 41(5), 801–819.
- Strohmer, J. & Mischo, C. (2015). The development of early childhood teachers' language knowledge in different educational tracks. *Journal of Education and Training Studies*, 3(2), 126–135.
- Stuck, A., Kammermeyer, G. & Roux, S. (2016). The reliability and structure of the classroom assessment scoring system in German pre-schools. *European Early Childhood Education Research Journal*, 24(6), 873–894.

- Swiss Society for Research in Education Working Group. (2010). *EDUG user guide*. Neuchatel: RDP.
- Tang, T. K. W. & Watkins, D. (1994). Towards a taxonomy of teacher thinking. *Educational Research Journal*, 9, 38–43.
- Textor, M. R. (2007). *Bildung in der Erzieherin-Kind-Beziehung*. Zugriff am 07.11.2019 unter <https://www.kindergartenpaedagogik.de/fachartikel/gruppenleitung-erzieherin-kind-beziehung-partizipation/beziehungsgestaltung-gespraechsfuehrung-konflikte/1119>.
- Tietze, W., Becker-Stoll, F., Bensel, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B., Keller, H. & Leyendecker, B. (Hrsg.). (2012). *NUBBEK. Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit. Fragestellungen und Ergebnisse im Überblick*. Zugriff am 07.11.2019 unter www.nubbek.de/media/pdf/NUBBEK%20Broschuere.pdf.
- Tietze, W. & Viernickel, S. (Hrsg.). (2007). *Pädagogische Qualität in Tageseinrichtungen für Kinder: Ein nationaler Kriterienkatalog*. Weinheim: Beltz.
- van Es, E. & Sherin, M. (2002). Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *Journal of Information Technology for Teacher Education*, 10(4), 571–596.
- van Someren, M., Banard, Y. & Sandberg, J. (1994). *The Think aloud Method: A practical guide to modelling cognitive processes*. London: Academic Press.
- von Suchodoletz, A., Fäsche, A., Gunzenhauser, C. & Hamre, B. K. (2014). A typical morning in preschool: Observations of teacher–child interactions in German preschools. *Early Childhood Research Quarterly*, 29(4), 509–519.
- Wadepohl, H. & Mackowiak, K. (2016). Beziehungsgestaltung und deren Bedeutung für die Unterstützung von kindlichen Lernprozessen im Freispiel. *Frühe Bildung*, 5, 22–30.
- Wasik, B. A. & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology*, 103(2), 455–469.
- Webb, N. & Shavelson, R. (2005). Generalizability theory: Overview. In B. Everitt & D. Howell (Hrsg.), *Encyclopedia of Statistics in Behavioral Science*. Hoboken, NJ: Wiley.
- Weltzien, D. (2014). *Interaktion als didaktisches Prinzip: Bildungsprozesse bewusst begleiten und gestalten*. Weinheim: Beltz Juventa.
- Weltzien, D., Fröhlich-Gildhoff, K., Wadepohl, H. & Mackowiak, K. (2017). Interaktionsgestaltung im familiären und frühpädagogischen Kontext: Einleitung. In H. Wadepohl, K. Mackowiak, K. Fröhlich-Gildhoff & D. Weltzien (Hrsg.), *Interaktionsgestaltung in Familie und Kindertagesbetreuung* (S. 1–26). Heidelberg: Springer.
- Whimbey, A., Vaughan, G. & Tatsuoka, M. (1967). Fixed effects vs. random effects: Estimating variance components from mean squares. *Perceptual and Motor Skills*, 25, 668.
- Wildgruber, A., Wertfein, M., Wirts, C., Kammermeier, M. & Danay, E. (2016). Situative Unterschiede der Interaktionsqualität im Verlauf des Kindergartenalltags. *Frühe Bildung*, 5, 206–213.
- Williams, E.H.J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research*, 2, 149–168.
- Wirts, C., Cordes, A.-K., Egert, F., Fischer, S., Radan, J., Reber, K., Reichl, S., Schauland, N., Quehenberger, J., Danay, E. & Becker-Stoll, F. (2018). *Zwischenbericht zu den formativen und summativen Evaluationen (BiSS-E1 und BiSS-E2) der Bund-Länder-Initiative Bildung durch Sprache und Schrift*. München: Staatsinstitut für Frühpädagogik.

Ina E. Rüber

Continuation and changes in civil participation during adulthood: A matter of education and learning?

Abstract

This paper revisits the associations between education and learning with civil participation during adulthood, focusing on the questions whether and how adult learning may exert an effect on civil participation. Adult learning is distinguished upon its degree of institutionalization. Using data from the German National Educational Panel Study (NEPS) further allows to differentiate recurring civil participation and the uptake of civil participation. Multinomial logistic regression models reveal that the odds of recurring civil participation, against the odds of no civil participation, increase with educational degree, non-formal and informal learning. This result supports earlier research. The models further show that the odds of taking up civil participation during adulthood, against the odds of no civil participation, do not change with educational degree but with informal learning and especially with non-formal learning. The results on the one hand, speak in favor of the continuity theory of aging; civil participation behavior establishes early in life and upon similar determinants as the dispositions for adult learning. On the other hand, non-formal learning may trigger adults to uptake civil participation by providing opportunities to extend social networks.

Keywords

Adult learning; Education; Adult education; Civil participation; Volunteering

Kontinuität und Wandel zivilgesellschaftlicher Partizipation im Erwachsenenalter: Eine Frage der Bildung und des Lernens?

Zusammenfassung

Dieser Artikel untersucht den Zusammenhang zwischen Bildung und Lernen mit zivilgesellschaftlicher Partizipation (zP) im Erwachsenenalter. Fokussiert wird die Fragestellung inwieweit das Lernen Erwachsener einen Effekt auf die

Ina Rüber, M. Sc., Wuesterather Hof 1, 56727 Mayen
e-mail: ina.rueber@gmail.com

zP hat. Es werden formales, non-formales und informelles Lernen unterschieden. Anhand der Daten des Nationalen Bildungspanels (NEPS) kann weiterhin zwischen wiederkehrender zP und der erstmaligen Aufnahme von zP unterschieden werden. Die Ergebnisse multinomialer logistischer Regressionsmodelle zeigen, dass wiederkehrende zP im Vergleich dazu nicht zu partizipieren mit Bildung, non-formalen und informellen Lernen wahrscheinlicher wird. Das zeigte bereits frühere Forschung. Die Modelle zeigen jedoch auch, dass die Aufnahme zP im Erwachsenenalter unabhängig von Bildung ist, aber mit informellen und insbesondere mit non-formalen Lernen wahrscheinlich wird. Einerseits stützen die Ergebnisse die Theorie der Kontinuität des Alterns; zP festigt sich bereits in der Jugend und im jungen Erwachsenenalter durch ähnliche Einflüsse, die auch die Dispositionen zum Lernen im Erwachsenenalter prägen. Andererseits kann die Aufnahme von zP teilweise durch non-formales Lernen erklärt werden, denn es ermöglicht die Erweiterung sozialer Netzwerke, die entscheidend für die Aufnahme zP sind.

Schlagworte

Lernen Erwachsener; Bildung; Erwachsenenbildung; Zivilgesellschaftliche Partizipation; Freiwilliges Engagement

1. Introduction

Civil participation includes regularly performed activities that serve the good of the society. It builds the fundamental base for a cohesive, integrative and functioning democratic society (Priemer, Krimmer, & Labigne, 2017; Putnam, 2000). Consequently, there is great political and public interest in achieving high civil participation rates, and there is great academic interest in explaining differences in individual civil participation behavior. In the academic discourse, one of the most robust findings is that there is a positive association between civil participation and education. The more time an individual spends in formal education and the higher the educational degree, the greater is his or her chance for civil participation (Gesthuizen, Van der Meer, & Scheepers, 2008; Mascherini, Vidoni, & Manca, 2011; Wilson, 2012). This is not limited to full time primary and higher education; also indeed, participation in adult learning positively relates to civil participation (Schuller, Preston, Hammond, Brassett-Grundy, & Bynner, 2004; Thomas, 2017; Vera-Toscano, Rodrigues, & Costa, 2017). Whether these associations, especially the one with adult learning (Field, 2011), are causal, to what extent, and at which stage of life they occur, are all, however, still matters for investigation. This article closely investigates the association of adult learning and civil participation.

In line with the continuity theory of normal aging, which implies that individuals are inclined towards psychological continuity and continuity of social behavior (Atchley, 1989; Wilson, 2012), for the majority of adults, civil participation behav-

ior does not change much over the life course (Lancee & Radl, 2014). Accordingly, civil participation during childhood and adolescence substantially increases the probability for recurring civil participation during adulthood (Oesterle, Johnson & Mortimer, 2004). The positive association between an adult's educational degree and his or her civil participation may, hence, only manifest due to an effect of education on civil participation early in life and the continuity of civil participation behavior thereafter (Schnittker & Behrman, 2012). The plausibility that further educational activities during adulthood, adult learning, exerts an effect on civil participation is also low, when civil participation tends to remain stable. However, if civil participation changes over the life-course, these changes may partly be triggered by adult learning.

Changes in civil participation behavior are functions of its changing determinants, such as the individuals' motivations, capacities and networks of recruitment (according to the *Civic Voluntarism Model* by Verba., Schlozman, & Brady, 1995) or the interplay of individuals' preferences, resources and external constraints (or opportunities) with those of public organizations (according to the volunteering and membership market model by Wiertz, 2016). Several studies assume key determinants of civil participation on the individual level to change with adult learning and, therefore, they expect an effect of adult learning on civil participation (for an overview, see Rüber, Rees, & Schmidt-Hertha, 2018). The empirical evidence for such an effect, however, is rather weak and even weaker when it comes to identifying the mechanisms at play. The majority of the studies either worked with cross-sectional data (e.g. Vera-Toscano, et al., 2017), investigated very specific cases (e.g. McIntyre, 2012) or analyzed a peculiar type of adult learning (e.g. Ruhose, Thomsen, & Weilage, 2019). As a consequence, they run into problems of unobserved heterogeneity, problems of weak generalizability of the results or the implications of the results for the potential mechanisms of an effect remain rather limited. The present article, aims at furthering the debate by adding an empirical analysis on a representative sample for Germany, which targets high generalizability, relatively low problems of unobserved heterogeneity and an advanced discussion of mechanisms. It theoretically follows the approach by Rüber et al. (2018), who argue that the mechanisms by which adult learning may exert an effect on civil participation depend on the degree of institutionalization of adult learning. Therefore in the analysis, I distinctively model the associations of formal, non-formal and informal adult learning with civil participation. To further the debate on causality in the association, I distinguish between recurring civil participation and the recent uptake of civil participation. The results reveal important insights on the potential mechanisms between adult learning and civil participation, which will be discussed further in the course of this article.

The next section entails a brief conceptualization of civil participation and adult learning and discusses the German context. The following section entails theories on the mechanisms behind the associations of informal, non-formal and formal adult learning with recurring civil participation and the uptake of civil participation. After a description of the data and the analytical strategy, I present the main

empirical results of the analysis. The remainder of the article entails a comprehensive discussion of the findings.

2. Civil participation and adult learning in the German context

2.1 Civil participation

From a political science point of view, civil participation is a latent form of political participation (e.g. Ekman & Amnå, 2012) and includes two dimensions: involvement (e.g. personal interest in politics) and civic participation (e.g. volunteering). Putnam (2000) uses the term in an even broader manner; civil participation includes any kind of individuals' cultural, social and political participation as well as values and attitudes towards the society. Both conceptualizations of the term include observable characteristics, such as actions, and unobservable characteristics, such as perceptions. In the center of interest of this article are the observable characteristics of civil participation, hence actions, and it further focusses only on collective forms of civil participation (e.g. volunteering), which are strongly embedded in the functioning of the German society.

Alscher, Droß, Priller, and Schmeißer (2013) refer to Germany as the country of clubs and civil organizations. There are more than 600,000 civil organizations in Germany and on average memberships are rising. At the same time, the number of individually organized and project-shaped civil activities has also significantly increased (Simonson & Vogel, 2017). According to data from the *German Survey on Volunteering*, the number of voluntarily engaged citizens rose from 35.5 percent in 1991 to 44.3 percent in 2014 (Vogel, Hagen, Simonson, & Tesch-Römer, 2017). 52 percent of highly educated respondents and 55 percent of the respondents who are still in education report on civil participation in 2014. The share among respondents with low education levels is significantly lower (28 %). Compared to earlier survey years, the differences in civil participation rates between educational groups increased, while the differences in civil participation rates between age groups decreased. The civil participation share among adults above the age of 65 is significantly lower, as compared to younger age groups. Still, civil organizations and associations are confronted with a rising average age amongst their members and they face difficulties in the recruitment of new and younger active members (Alscher et al., 2013).

A study by Lancee and Radl (2014) revealed that in Germany, the majority of individuals show high continuity in their civil participation behavior over the life course. Using data from the German *Socio-Economic Panel* (SOEP), the authors demonstrate that about two thirds of the respondents do not change their volunteering behavior during 25 years of observation. Half of the respondents never volunteered, while 15 percent of the male respondents and 10 percent of the female

respondents volunteered throughout. A third of the respondents change their volunteering behavior, by either taking up volunteering or by ceasing to volunteer during adulthood.

2.2 Adult learning

Following earlier research, I differentiate adult learning, the learning of adults after completing initial full-time education and entering the labor market, into formal, non-formal and informal learning (Buchholz, Unfried, & Blossfeld, 2014; Eisermann, Janik, & Kruppe, 2014; Eurostat, 2017). Formal adult learning activities are the most institutionalized learning activities. They culminate in the acquirement of a nationally accredited degree and therefore follow pre-established curricula. Non-formal adult learning does not lead up to a nationally accredited degree, but also incorporates pre-established curricula. It encompasses adult education courses and further training. Finally, informal learning is the least institutionalized type of adult learning. It is not bound to a curriculum, or to certain places or times. It comprises any individual and independent knowledge acquisition.

The *Adult Education Survey 2016* (AES) provides data on adult learning activities from a representative sample of the German residential population (age 18-69). The respondents report on their participation in learning during the time period of twelve months prior to the interview. The data reveals that formal learning is the least dominant type of adult learning, with a participation rate of 10 percent (Kuper, Christ, & Schrader, 2017). Formal adult learning, here, also includes learning that the respondents would describe as part of their initial education. The percentage drops to about 3 percent, when counting only those adults, who regard their formal learning as further education. Half of the AES respondents between the ages of 18 and 64 took part in non-formal learning activities (Bilger & Strauß, 2017) and 43 percent engaged in informal learning activities (Kaufmann-Kuchta & Kuper, 2017). Informal learning here comprises acquiring knowledge from friends, family or colleagues, from books, by searching online or using other media, or by visiting museums or libraries. Eisermann and colleagues (2014) show that adult learning participation rates differ substantially between different data sources because of measurement heterogeneities. Accordingly, using NEPS data, Kruppe and Trepesch (2017) found much higher informal learning participation rates compared to the AES. As such, 63 percent of the NEPS-respondents took part in informal learning within a time period of 12 months (2012/13), while informal learning, here, comprised reading specialized literature, attending lectures or congresses or using self-learning programs. The figures that Kruppe and Trepesch (2017) show for formal learning do not refer to a time period of 12 months, but they refer to the entire life course. Naturally, the reported formal learning rate of 39 percent is higher as compared to the AES formal learning rate. The share of NEPS respondents, who engage in non-formal learning, is a little lower (around 40 %) compared to the AES figures.

According to AES data, 60 percent (57 %) of adults who are highly educated participate in non-formal (*informal*) learning, while only 35 percent (32 %) of low educated adults do. Similarly, large differences occur when comparing non-formal learning participation between employed (56 %) and unemployed (27 %) adults (Bilger & Strauß, 2017). These figures correspond to the observation that in 2016, 71 percent of non-formal adult learning in Germany was in-service training and another 10 percent was still at least work-related. Even participants' purposes for engaging in informal adult learning are principally job-related (Rüber & Bol, 2017).

3. Mechanisms behind the associations of learning and civil participation during adulthood

As touched upon in the introduction, there is a rich set of empirical work that reports a positive association between adult learning and a variety of civil participation measures, such as volunteering or active club memberships (Bynner & Hammond, 2004; Feinstein, Hammond, Woods, Preston, & Bynner, 2003; Preston & Feinstein, 2004; Ruhose et al., 2019; Schuller & Desjardin, 2010). Rarely, however, these civil participation measures capture change. They often represent a snapshot of whether or not respondents take part during a certain period of time related to the interview. Those reporting on civil participation may have recently started civil participation or they may report on recurring civil participation. To reach a better understanding on mechanisms behind the association of adult learning and civil participation, I separately discuss recurring civil participation and the uptake of civil participation.

3.1 Adult learning and recurring civil participation

According to the continuity theory of aging (Atchley, 1989; Wilson, 2012), recurring civil participation during adulthood results from learned behavior in earlier phases of the life course. This is in line with research on the determinants of civil participation, which shows that the strongest predictors of civil participation are those, which tend to remain rather stable over the life course, such as education (Nie, Junn, & Stehlik-Barry, 1996; Wilson, 2012). Further, Verba and colleagues (1995) state that socialization processes in family and school lay the foundations for the main determinants of civil participation: individuals' motivations, capacities and networks of recruitment. In regards to the mechanisms behind the association between adult learning and recurring civil participation, this implies that either the association only comes about due to common stable determinants (1) or because recurring civil participation opens up opportunities for adult learning (2).

(1) Similarly to civil participation, especially non-formal and informal adult learning are more common among highly educated adults (Kruppe & Trepesch,

2017). Positive experiences with education and learning in early life further increase the chances of education and learning later in life (Gorges, 2018). Besides initial education, other common determinants of civil participation and adult learning are personality traits, such as prosocial behavior, intelligence and general curiosity (Van Ingen & Dekker, 2011). (2) Civil participation often includes possibilities for non-formal and informal adult learning. For example, clubs or associations often offer trainings to their volunteers, to increase their motivations and performances (Simonson & Gordo, 2017). Moreover, volunteers, who want their actions to have a positive impact, often engage in informal learning to increase their efficacy (Brödel, 2006).

3.2 Adult learning and the uptake of civil participation

Taking up civil participation during adulthood means to change civil participation behavior and therewith to break with continuity. However, the same theoretical models that explain recurring civil participation, may also explain changes in civil participation. Even if the foundations of motivations, capabilities and networks of recruitment are laid early in life, they may still become subject of change over the life course. Earlier research for instance reveals that life-course transitions related to the family domain can trigger such changes (Janoski & Wilson, 1995). Lancee and Radl (2014) show on the basis of fixed effects models with German panel data that having pre-school children decreases parents' volunteering, while having children in school increases parents' volunteering. Here, the capabilities in terms of free time available for civil participation change dramatically. Marriage and divorce also showed to decrease volunteering. Here, both motivations but also networks of recruitment possibly change. For my research question, this, in connection with the earlier considerations, implies that there are three possible explanations for a positive association between adult learning and the uptake of civil participation.

First, adult learning can be a requirement for the uptake of civil participation. For example, volunteering for ambulance services is only possible if the volunteer took part in some preparatory training. This proposition, however, does not apply to formal and informal adult learning and therewith only suggests an association between non-formal adult learning and the uptake of civil participation. Second, certain life-course transitions or events may affect both participation in adult learning and the uptake of civil participation. Here, one important factor is time, which is both required for adult learning (Mania, 2019) and civil participation. The positive effect of children turning into school age on civil participation for parents, equally holds for participation in adult learning. The surplus of free time can be invested especially in informal learning activities, which are not bound to a certain schedules. Third, adult learning may trigger the uptake of civil participation by changing motivations, capabilities or networks of recruitment for civil participation (Verba et al., 1995; or resources, beliefs and external constraints as of Wiertz, 2016).

Reviewing studies on returns to adult learning and theoretical approaches on determinants of civil participation, Rüber et al. (2018) provide a theoretical framework, which summarizes five potential mechanisms of an effect of adult learning on the uptake of civil participation. First, successful participation in formal adult learning and, occasionally, participation in non-formal adult learning, improve individuals' *economic conditions*. The obtained qualifications can result in a wage increase, a promotion or a change of employer (Ehlert, 2017; Hanushek, Schwerdt, Woessmann, & Zhang, 2017). Economic conditions are comparable to individuals' capabilities or resources for civil participation (Verba et al. 1995; Wiertz, 2016). Civil participation requires investing time, which is unpaid. The investment of unpaid time may only be possible and desirable if there are no private economic shortages. Second, formal and non-formal adult learning often take place within social groups. Depending on the teaching style and the focus of the course, the learner is able to expand his or her *social networks* (networks of recruitment, opportunities). The greater and the more heterogenic the number of social contacts, the likelier it is that the participant will meet people who are engaged in civil participation. Contact with other civil participants increases the chance of being asked to engage in civil participation. Being actively recruited has strong effects on the decision and the opportunities to engage in civil participation (Einolf & Chambré, 2011; Hustinx, Cnaan, & Handy, 2010). Third, the new knowledge acquired via adult learning can alter one's *attitudes and values* towards civil participation and therewith motivations for civil participation or beliefs on civil participation may change. This applies to all types of adult learning. A simple example is adult learning about environmental issues. If the learner did not know about issues like global warming or ocean pollution before, the learning activity is likely to trigger a change in the learner's attitudes towards consumption behavior, and may trigger engagement within environment protection. Fourth, if the learning activity connects to a sphere of civil participation, as in the given example, it may not only change the learners' attitudes and values. It may also facilitate the learner with new *qualifications* (capabilities or resources) for civil participation. Owning these qualifications in turn can increase the motivation to uptake civil participation. Finally, positive learning experiences in every context may change adults' *low-level personality trait characteristics* (motivations, capabilities, resources). It can foster the individual's self-efficacy and self-confidence. These character traits are likely to increase the chance of starting new activities such as civil participation.

In regards to a potential effect of adult learning on the uptake of civil participation, Rüber et al. (2018) suggest that the mechanisms related to personality, values, or knowledge and skills may apply to all types of adult learning. Mechanisms related to networks and communities call for a social setting, which does not occur in informal learning. Economic conditions change most significantly with formal adult learning. Hence, if all types of adult learning show the same association with the uptake of civil participation, only those mechanisms, which work equally for all types are plausible. Respectively, if the association only shows for specific types of adult learning, only the respective mechanisms may apply.

Rüber et al. (2018) acknowledge that next to the degree of institutionalization, the status and the content of adult learning are relevant as to whether (and to what extent) they might affect civil participation. In this analysis, however, I only focus on the degree of institutionalization.

To sum up the theoretical considerations of the article, there are three ways by which the association between adult learning and civil participation may come about: common determinants, an effect of civil participation on adult learning and an effect of adult learning on civil participation. My focus is on the latter, which I conjecture to be only plausible for the uptake of civil participation and to vary between formal, non-formal and informal adult learning. An effect of civil participation on formal adult learning is not expected.

4. Data

To investigate the associations of adult learning and civil participation, I use data from the German *National Educational Panel Study* (NEPS) (Blossfeld, Roßbach, & von Maurice, 2011). The NEPS collects longitudinal data from adults within its sixth starting cohort *Adult Education and Lifelong Learning* (SC6) via computer-assisted personal interviewing and telephone interviewing (CAPI and CATI). While it meets the aim of capturing learning activities within different contexts across the whole life span, it is not equally able to capture civil participation. Data on civil participation was, so far¹, only captured in the fifth NEPS-main survey wave (2013/2014) out of eight available waves. The models in this article, therefore, mainly build on a cross-sectional sample. However, since the NEPS is a panel study it is possible to incorporate control variables related to life-transitions that are shown to influence volunteering during adulthood and to run robustness checks with lagged predictors. This improves the chances of eliminating problems of unobserved heterogeneity and reversed-causality.

The original sample of the fifth NEPS wave ($N = 10,639$) consists of three subsamples. The first subsample builds on the predecessor study *Working and Learning in a Changing World* (ALWA 2007/2008) (Kleinert et al., 2011). The second subsample is a refreshment sample and a pile up sample drawn from the first NEPS wave (2009/2010). The third subsample is a refreshment sample in the third NEPS wave (2011/2012). Using a two-stage cluster sampling approach with municipalities as first sampling unit and persons as second sampling unit, the NEPS aims to achieve a representative sample for the German working-age population (cohorts 1944–1986) in each wave (Aßmann et al., 2011). Next to the refreshments that combat panel attrition, the data also entails weights that are calculated using *Mikrozensus* data. Each analysis presented in this paper incorporates both the design weights and the post-stratification weights.

1 doi:10.5157/NEPS:SC6:8.o.o

5. Variables

Within the scope of this paper, I discuss the associations between formal, non-formal and informal learning and civil participation, focusing on potential effects of adult learning on civil participation. I measure civil participation with a categorical variable: (1) no civil participation, (2) recurring civil participation, (3) uptake of civil participation. It builds on two survey questions. The first question asks whether the respondent “has been actively involved in one or more areas since the last interview”. Four areas² are mentioned to the respondent successively and the respondent is requested to indicate whether he or she was actively involved. It is not possible to accurately demarcate in which area the respondent was involved from the data, which leads to a very broad measurement of civil participation. To distinguish between recurring civil participation and recent uptake of civil participation, I use additional information from the follow-up question, which asks whether the respondent has “ever been actively involved in clubs, organizations or self-help groups before”. If the respondent indicates having been involved before, I rate the civil participation as recurring civil participation. If the respondent indicates not having been involved in civil participation in the past, I assume that the respondent took up civil participation between the current and the latest interview date. To ensure that the uptake of civil participation took place in this period, I delete the cases that report on civil participation, but that temporarily dropped out of the survey in the previous wave. This step leads to losing 101 observations, but it is a necessary step, since the predecessor question asks for participation between the last interview and the current date and that period differs strongly in the case of a temporary drop out.

I generate four dummies for the independent variables; one variable each for formal and non-formal and two variables for informal adult learning. The formal learning dummy takes on the value 1 if the respondent reported on at least one education spell³ in the twelve months period prior to the interview (after having left initial full-time education for a minimum of 12 months). Using the *Further Education* data file, the non-formal learning dummy takes on the value 1 if the respondent took part in at least one course in the 12 months prior to the interview. Finally, I use two measures for informal learning, out of four measures that are provided in the NEPS data. The measures that are not included in this analysis are strongly tied to the world of work and have a weaker match to earlier research on individual informal learning. The selected measures (see Table 1) correspond to the measures used in the AES.

2 (1) clubs or organizations, including sports clubs, political parties, trade unions and church communities, (2) voluntary fire services and rescue services, carnival associations and welfare organizations, (3) citizens’ initiatives and citizens’ clubs, parents’ associations, theatre and music groups or initiatives for the unemployed, and (4) work as a lay judge, visiting the sick or participation in solidarity projects.

3 sptype = School, VocPrep or VocTrain in Biography data file.

Table 1: Weighted descriptive statistics of dependent, independent and control variables

| | <i>M</i> | <i>SD</i> | Range | Description / Survey Question |
|---------------------------------|----------|-----------|----------|---|
| <i>Civil participation (cp)</i> | | | | |
| No cp | .53 | .50 | 0/1 | (1) No civil participation since the last interview (0) else |
| Recurring cp | .27 | .45 | 0/1 | (1) Civil participation since the last interview and before (0) else |
| Uptake of cp | .19 | .40 | 0/1 | (1) Civil participation since the last interview but not before (0) else |
| <i>Adult Learning</i> | | | | |
| Formal | .07 | .25 | 0/1 | (1) Yes (0) No |
| Non-formal | .48 | .50 | 0/1 | (1) Yes (0) No |
| Informal: Reading | .57 | .49 | 0/1 | <i>Since the last interview, did you read textbooks and specialized books or professional magazines, to learn more in the professional or private field?</i> (1) Yes (0) No |
| Informal: Media | .23 | .42 | 0/1 | <i>Have you used any computerized learning programs, learning CDs or DVDs or similar materials since the last interview in order to enhance your private or professional knowledge?</i> (1) Yes (0) No |
| <i>Controls</i> | | | | |
| <i>Education</i> | | | | |
| Low | .27 | .44 | 0/1 | (1) no or elementary education (0) else |
| Medium | .53 | .50 | 0/1 | (1) intermediate and general/vocational maturity (0) else |
| High | .21 | .40 | 0/1 | (1) lower and higher tertiary education (0) else |
| Gender | .50 | .50 | 0/1 | (1) Female (0) Male |
| <i>Migration status</i> | | | | |
| Native | .75 | .43 | 0/1 | |
| 1st generation | .17 | .38 | 0/1 | |
| 2nd generation | .08 | .27 | 0/1 | |
| <i>Parents' education</i> | | | | |
| None | .83 | .38 | 0/1 | (1) No parent has tertiary education (0) else |
| One | .12 | .33 | 0/1 | (1) One parent has tertiary education (0) else |
| Both | .05 | .21 | 0/1 | (1) Both parents have tertiary education (0) else |
| <i>Health</i> | | | | |
| (Very) good | .67 | .47 | 0/1 | <i>I now have a brief question about your health. How would you generally describe your state of health?</i> |
| Average | .28 | .45 | 0/1 | |
| (Very) poor | .05 | .23 | 0/1 | |
| Age | 47.39 | 11.50 | 26-69 | Mean centered in the analysis |
| Children in HH | .45 | .50 | 0/1 | (1) at least one child lives in the household (0) no child |
| Household income | 2058.70 | 877.94 | 849-6010 | Net household income divided by square root of the household size; mean centered in the analysis |
| East | .20 | .40 | 0/1 | Residence in (1) East Germany (2) West Germany |
| <i>Transitions</i> | | | | |
| Div./wid. | .01 | .07 | 0/1 | (1) divorced or widowed in past 12 months (0) no |
| Married | .02 | .13 | 0/1 | (1) married in past 12 months (0) no |
| Parent | .02 | .13 | 0/1 | (1) started living with child in past 12 months (0) no |

Note. *N* = 8317. Source: Own calculations using 5th orig. NEPS wave (SC6), version 8.0.0, doi:10.5157/NEPS:SC6:8.0.0.

To account for theoretical propositions that adult learning and civil participation have common determinants, I control for variables that earlier research showed to influence both adult learning and civil participation. These include both stable variables; the educational level, sex, migration background and parental education, and variables that are subject to change during the life-course; presence of children in the household, household income, residency in East or West Germany and health condition. I also compute variables to capture recent life-course transitions within the 12 months prior to the interview, such as getting a divorce or being widowed, having a child or getting married. Table 1 includes the descriptive statistics for all variables. The analytic sample shrinks down to 8,316 respondents, since I exclude all missing values. The inclusion of the household income leads to the greatest loss of cases, but is worth including since research showed its potential connection with civil participation (Mascherini et al., 2011).

6. Analytical strategy

The analytical strategy which underlies the empirical analyses presented in this paper, targets at disentangling which mechanisms of the ones described in the theory section apply in regards to the associations between adult learning and civil participation. Following the theoretical approach, civil participation is measured using a nominal categorical variable, distinguishing between recurring-, uptake of- and no civil participation. This calls for an estimation technique that accounts for the nonlinearity and disordered nature of the outcome categories. Therefore, I employ multinomial logistic regression models. Multinomial logistic regressions are effectively an extension of binary logistic regressions, while modeling each pair of the outcome categories simultaneously (Agresti & Finlay, 2009, pp. 51ff.; Treiman, 2009, pp. 336ff.). It predicts the odds of being in one category of the outcome variable against a baseline category. In my models I define the baseline category as being in the group with no civil participation. Hence, I estimate the odds of respondents falling into the group of recurring civil participation or uptake of civil participation against falling into the group of no participation.

To disentangle the different explanations, I apply a stepwise approach. The first model solely estimates the associations of the different types of adult learning with civil participation. The second model adds initial education as the main common determinant of civil participation and adult learning. The third model incorporates all control variables.

7. Results

Prior to the multivariate results, some descriptive statistics already provide insights into the major contribution of this article.

7.1 Descriptive results

Table 2 provides an overview of civil participation for initial education and adult learning. In general, the largest group of respondents do not report on civil participation (59 %), while 23 percent are recurring participants and 18 percent took up civil participation in the year prior to the survey interview. These figures correspond to the figures from the German Survey on Volunteering (*FWS 2014*), which reports a civil participation rate of 44 percent (Simonson, Hameister, & Vogel, 2017). This is a good indicator for the accuracy and comparability of the NEPS civil participation measure, sampling and weighting strategy.

Table 2: Weighted proportions of civil participation by education, formal, non-formal and informal adult learning

| | Civil Participation | | |
|---|---------------------|-------------|-------------|
| | No | Recurring | Uptake |
| Education | | | |
| Low | .64 | .18 | .19 |
| Medium | .59 | .22 | .18 |
| High | .52 | .32 | .16 |
| Formal adult learning | | | |
| No | .59 | .23 | .18 |
| Yes | .61 | .21 | .18 |
| Non-formal adult learning | | | |
| No | .64 | .19 | .16 |
| Yes | .53 | .27 | .20 |
| Informal adult learning: Reading books and specific literature | | | |
| No | .68 | .15 | .17 |
| Yes | .52 | .29 | .19 |
| Informal adult learning: Using digital media for knowledge acquisition | | | |
| No | .61 | .21 | .18 |
| Yes | .54 | .29 | .18 |
| Total | .59 | .23 | .18 |
| Number of observations | 8317 | 8317 | 8317 |

Note. Source: Own calculations using 5th orig. NEPS wave (SC6), version 8.0.0, doi:10.5157/NEPS:SC6:8.0.0.

As expected from earlier research, the share estimate for recurring civil participation is highest for highly educated respondents (32 %) and substantially lower for those with medium (22 %) and low education (18 %). This pattern, however, does not hold for those, who took up civil participation. The share estimates are comparatively similar between the educational groups with opposing tendencies as compared to recurring civil participation. 19 percent of the low-educated, 18 percent of the medium educated and 16 percent of the highly educated took up civil participation. The relatively low share of highly educated adults in this group is likely to result from the high share in the group of recurring civil participation.

There are no substantial differences in civil participation behavior between respondents, who took part in formal learning and those, who did not. Large differences in civil participation behavior exist in regards to non-formal learning. 27 percent of non-formal learners report recurring civil participation, while only 19 percent of those, who do not take part in non-formal learning, report on recurring civil participation. Comparable figures show for the uptake of civil participation (20 % vs. 16 %). Table 2 further reveals that the measures chosen for informal learning show differing associations with civil participation. While the acquirement of knowledge by reading books and specific literature shows similar patterns as those described for non-formal learning, using new media for knowledge acquisition is unrelated to taking up civil participation, but occurs more often in relation to recurring civil participation.

7.2 Multivariate results

Table 3 reports the results from the multinomial logistic regression models. Model 1 shows significant associations of adult learning and recurring civil participation for formal, non-formal and informal learning by reading books or specific literature. Informal learning using digital media does not significantly associate. Non-formal and informal learning by reading increase the odds of recurring civil participation against no civil participation, while formal learning decreases the odds. The strongest association shows with informal learning by reading. The odds to uptake civil participation against no civil participation do not associate with formal and informal learning by using digital media, but the odds increase with non-formal learning and informal learning by reading.

Adding the highest educational degree to the model leads to significant changes in the associations between non-formal learning and informal learning by reading on recurring civil participation. The effect sizes both decrease significantly. There is no such change in the other associations. Concerning education, model 2 reveals that higher education positively associates with the odds of being in the group of recurring civil participation rather than in the group of no civil participation. A higher educational degree however does not increase the odds of taking up civil participation against no civil participation. These findings also hold for model 3.

Table 3: Weighted multinomial logistic regressions of civil participation

| | Recurring civil participation | | | Uptake of civil participation | | |
|---|-------------------------------|---------------------|---------------------|-------------------------------|---------------------|---------------------|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| Formal learning | -0.438* (0.17) | -0.455** (0.17) | -0.235 (0.18) | -0.231 (0.19) | -0.223 (0.19) | -0.075 (0.19) |
| Non-formal learning | 0.428*** (0.08) | 0.396*** (0.09) | 0.332*** (0.09) | 0.348*** (0.08) | 0.362*** (0.09) | 0.377*** (0.09) |
| Informal: Reading | 0.832*** (0.09) | 0.761*** (0.09) | 0.755*** (0.09) | 0.344** (0.10) | 0.373*** (0.10) | 0.390*** (0.11) |
| Informal: Media | 0.165 (0.09) | 0.163 (0.10) | 0.199* (0.10) | -0.035 (0.11) | -0.034 (0.11) | -0.014 (0.11) |
| Education (<i>Ref. = Low</i>) | | | | | | |
| Medium | | 0.117 (0.12) | 0.185 (0.13) | | -0.048 (0.12) | 0.015 (0.13) |
| High | | 0.411** (0.13) | 0.407** (0.14) | | -0.185 (0.13) | -0.039 (0.14) |
| Age | | | 0.016** (0.00) | | | 0.013** (0.01) |
| Sex | | | -0.160 (0.09) | | | 0.016 (0.10) |
| Migration (<i>Ref. = Native</i>) | | | | | | |
| 1st generation | | | -1.094*** (0.16) | | | -0.299 (0.16) |
| 2nd generation | | | -0.333** (0.13) | | | -0.059 (0.14) |
| Child in HH | | | 0.624*** (0.08) | | | 0.411*** (0.09) |
| Health (<i>Ref. = Good</i>) | | | | | | |
| Medium | | | -0.052 (0.10) | | | -0.157 (0.11) |
| Poor | | | -0.315 (0.19) | | | -0.643** (0.20) |
| Household income | | | 0.000*** (0.00) | | | -0.000 (0.00) |
| Employment (<i>Ref. = Not employed</i>) | | | | | | |
| Part-time | | | -0.125 (0.13) | | | -0.017 (0.14) |
| Full-time | | | -0.365** (0.13) | | | -0.161 (0.15) |
| East | | | -0.397*** (0.10) | | | -0.069 (0.10) |
| Parents education (<i>Ref. = None</i>) | | | | | | |
| One | | | 0.034 (0.12) | | | -0.091 (0.13) |
| Both | | | 0.084 (0.21) | | | -0.600** (0.22) |
| Transitions | | | | | | |
| Divorced / Widowed | | | 1.206 (0.76) | | | -0.213 (0.61) |
| Birth | | | -0.540* (0.27) | | | 0.523 (0.30) |
| Married | | | 0.309 (0.36) | | | -0.625* (0.30) |
| Constant | -1.683*** (0.09) | -1.777*** (0.11) | -1.525*** (0.18) | -1.523*** (0.08) | -1.487*** (0.11) | -1.505*** (0.18) |
| <i>N</i> | 8317 | 8317 | 8317 | 8317 | 8317 | 8317 |

Note. Reference category is no civil participation. Multinomial logistic regression coefficients are displayed with standard errors in parentheses. Source: Own calculations using 5th orig. NEPS wave (SC6), version 8.0.0, doi:10.5157/NEPS:SC6:8.0.0.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Model 3 includes all control variables. Here, the association of formal adult learning and recurring civil participation loses significance, while the one with informal learning by media gains significance. The model unfolds that, controlling for major common determinants of adult learning and civil participation, non-formal learning increases the odds of recurring civil participation by 39 percent ($e^{0.332}$) and it increases the odds of taking up civil participation by 46 percent ($e^{0.377}$) as compared to no civil participation. The increase of the odds is even higher for informal learning by reading. Further, all else being equal, the odds to uptake civil participation for non-formal learning against the odds to be in recurring civil participation for non-formal learning are 5 percentage points higher ($e^{(0.377-0.332)}$). In the case of informal learning, the odds decrease by 70 percentage points ($e^{(0.390-0.755)}$).

A glance at the coefficients of the control variables manifests the picture that the determinants of recurring civil participations and the determinants of taking up civil participation differ quite substantially. The model also replicates earlier findings, showing a strong negative association of marriage and taking up civil participation and a positive effect of children in the household on taking up civil participation (cf. Lancee & Radl, 2014; Rotolo, 2000). A migration background shows a negative influence on recurring civil participation, but not on taking up civil participation (cf. Wiertz, 2016). Quite comprehensibly, poor health significantly decreases the odds of taking up civil participation, while it does not significantly associate with recurring civil participation.

There are three major conclusions that can be drawn from the results. First, the explanatory power of the variables differs substantially between estimating recurring civil participation and taking up civil participation. Second, formal adult learning and media supported informal adult learning are not associated with taking up civil participation. Third, non-formal adult learning and informal learning by reading books and specific literature are likely to contribute to the uptake of civil participation, although adult, who engage in informal learning are much more likely to be in the group of recurring civil participation against taking up civil participation.

7.3 Robustness checks

The analyses presented above have a couple of limitations, when it comes to their potential for causal inference. Some of them can be tackled by separately employing different estimation techniques. In this section, I present a summary of the main findings from two robustness checks. A detailed description may be drawn from the supplementary material.

First, the measures used for civil participation and adult learning refer to the same time period. This means that the main model may not differentiate between an effect of civil participation on adult learning and an effect of adult learning on civil participation. To combat this problem of reversed causality to a certain extent, I make use of the NEPS panel structure running two wave multinomial logistic re-

gressions. Therein, the fourth NEPS main wave provides the data for adult learning activities. This ensures that the learning activity precedes the uptake of civil participation. Apart from one major exception and slight decreases in the size of the coefficients, the results are similar to the results from the main model. The major exception is, that the association of informal learning with taking up civil participation loses significance. This supports the implication that non-formal learning contributes strongest to the chances of taking up civil participation.

Second, I run the models separately for each adult learning activity to avoid collinearity. The results are similar to modelling them all together.

8. Conclusion and discussion

This article revisited the associations between education and learning with civil participation during adulthood, focusing on the questions whether and how adult learning may exert an effect on civil participation. The political faith in such an effect is strong (Dohmen, 1998; European Commission, 2002; United Nations General Assembly, 2015), as is the empirical evidence for a positive association, but not particularly for an effect. Viewing civil participation behavior from a life course perspective, even raises considerable doubts on the plausibility of an effect, since civil participation often establishes early in life and then remains rather stable. Hence, I distinguished recurring civil participation from the uptake of civil participation during adulthood. In regards to the association of civil participation and education the results support that the continuity theory of aging applies to civil participation. Education positively associates with recurring civil participation, while the uptake of civil participation during adulthood does not relate to education.

To analyze the associations of adult learning and civil participation, following earlier work, I distinguished formal, non-formal and informal adult learning. I draw 4 main conclusions from the empirical analyses, which allow for drawing inferences about the mechanisms underlying the associations between adult learning and civil participation.

First, in every model specification formal adult learning neither significantly associates with recurring civil participation, nor with the uptake of civil participation. In line with the theoretically derived expectation, this implies that there is no effect of civil participation on formal adult learning. However, against the expectations, there are also no common determinants of civil participation and formal adult learning and there is no effect of formal adult learning on the uptake of civil participation. The absence of an effect of formal adult learning on the uptake of civil participation indicates that the proposed mechanism by which adult learning may exert an effect on civil participation and which only applies to formal learning, does not show. The *economic conditions* of adults were expected to change with formal learning and to influence the capabilities or resources for civil participation. The results indicate, they do not. Further, those mechanisms, which were expected

to apply equally to all types of adult learning, either do not apply to formal learning, or are too weak to actually trigger changes in civil participation.

The implications regarding the association of formal adult learning and civil participation are subject to some important limitations of the analysis. The described mechanism for an effect of formal adult learning on the uptake of civil participation may only set in after the completion of formal adult learning, which is not specified as such in the analysis. Ongoing formal adult learning is often very time consuming and it poses a considerable additional financial expenditure (Kuper et al., 2017). It is quite likely that adults engaging in formal learning rather lack resources to continue or to uptake civil participation. Focusing on completed accredited courses, Feinstein et al. (2003) did find a positive association with civil participation.

Second, non-formal adult learning positively associates with both recurring civil participation and the uptake of civil participation in every model specification, while the size of the log odds decreases with the restrictiveness of the models. Hence, all explanations given for the positive associations may hold; there are common determinants and there are likely to be effects in both directions. In line with earlier research, civil participation, whether timely or planned, can increase the chance for participation in non-formal learning activities. Non-formal learning can, through the suggested mechanisms, increase the chances for adults to continue or to uptake civil participation, while the evidence for the latter is stronger, also considering that the vast majority of non-formal learning activities are job-related.

Third, the association of informal learning and civil participation differs upon the type of informal learning. Informal learning by using digital media does not significantly associate with civil participation, suggesting no common determinants and no effects. Informal learning by reading books and specific literature shows strong positive associations with recurring civil participation, which weaken with the introduction of controls, and moderate positive associations with the uptake of civil participation, which strengthen with the introduction of controls and become insignificant in the two-wave model. These results hint towards common determinants and towards civil participation to raise opportunities for informal learning by reading books and specific literature. The likelihood of an effect of this type of informal learning on the uptake of civil participation is rather low. The results on informal learning, again, point towards the interpretation that proposed mechanisms that shall apply to all types of adult learning may either be too weak to exert an effect on civil participation or the theoretical considerations may not hold.

Fourth, synthesizing all findings and theoretical implications on the potential mechanisms of an effect of adult learning on the uptake of civil participation suggests that only those mechanisms that apply to non-formal learning may hold. These refer to *social networks*, which are argued to have high explanatory power for civil participation behavior (Verba et al., 1995; Wiertz, 2016; Wilson, 2012). Non-formal adult learning takes place within social settings and therewith often opens up opportunities for the participants' to expand their social networks. This increases the chance to get in touch with civilly active people, and simultaneous-

ly it increases the chances to be recruited or to develop an interest in civil participation.

With this conclusion, I do not intend to reduce non-formal adult learning simply to a place of social encounters. The analysis does not stand against theories which imply that adult learning generates knowledge, skills or self-perceptions that may lead someone into civil participation. It only confirms that on the aggregate level and irrespective of the content, there is still an effect emerging from a different causal path. This also implies that more specified types of informal or formal adult learning might increase the chance to uptake civil participation. Further, the positive effect through social mechanisms is likely to become stronger or weaker when controlling for other course characteristics. Assuredly, the next step is to distinguish between more nuanced types of non-formal adult learning.

This also bridges over to the remaining limitations of this article. The study focusses on the degree of institutionalization of adult learning, but disregards the content, how it is financed and how long the learning activities lasts. It is likely that the learner enters and leaves the learning activity differently depending on whether it was self- or externally financed. Similarly, a two-week-course probably differs from a two-day course in its opportunities to connect with other course participants. On the other hand, if it is about the content more than about intensity or institutionalization, it might be sufficient to read just one journal entry to be motivated to uptake civil engagement. This might apply or should apply in particular to the learning opportunities connected to civic education. Another limitation to the present study is that it is unable to fully identify the direction of a potential causal effect and it is still subject to unobserved variable bias. Moreover, the analysis does not include interactions with education, gender or age groups. I strongly suggest that future research looks closely into the effect of non-formal adult learning including those interactions, since there is evidence that the effect may differ between these groups (e.g. Van Ingen & van der Meer, 2011). Finally, future research may facilitate data which allows for differentiation between civil participation activities, such as active membership versus volunteering. It may further determine whether the uptake of civil participation lasts or whether it is only short term.

In spite of its limitations, this study points out the important but insufficiently addressed differences in explaining recurring civil participation and the uptake of civil participation during adulthood in connection with education and learning.

References

- Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences*. 4th edition. London: Pearson.
- Alscher, M., Droß, P. J., Priller, E., & Schmeißer, C. (2013). *Vereine an den Grenzen der Belastbarkeit* (No. 07). WZBrief Zivilengagement.
- Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., Rohwer, G., Rässler, S., & Blossfeld, H. P. (2011). 4 Sampling designs of the

- National Educational Panel Study: Challenges and solutions. *Zeitschrift für Erziehungswissenschaft*, 14(2), 51.
- Atchley, R. C. (1989). A continuity theory of normal aging. *The gerontologist*, 29(2), 183–190.
- Bilger, F., & Strauß, A. (2017). Beteiligung an non-formaler Weiterbildung. In F. Bilger, F. Behringer, H. Kuper, & S. Schrader (Eds.), *Weiterbildungsverhalten in Deutschland 2016 – Ergebnisse des Adult Education Survey (AES)* (pp. 25–55). Bielefeld: wbv Media.
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von (Eds.). (2011). *Education as a life-long process. The German National Educational Panel Study (NEPS)*. Wiesbaden, Germany: VS.
- Brödel, R. (2006). Bürgerschaftliches Engagement und Weiterbildung. *Report. Zeitschrift für Weiterbildungsforschung*, 3(29), 70–80.
- Buchholz, S., Unfried, J., & Blossfeld, H.-P. (2014). Reinforcing social inequalities? Adult learning and returns to adult learning in Germany. In H.-P. Blossfeld, E. Kilpi-jakonen, D. Vono de vilhena, & S. Buchholz (Eds.), *Adult learning in modern societies: an international comparison from a life-course perspective* (pp. 242–263). Cheltenham: Edward Elgar Publishing.
- Bynner, J., & Hammond, C. (2004). The benefits of adult learning: quantitative insights. In T. Schuller, J. Preston, C. Hammond, A. Brassett-Grundy and J. Bynner (Eds.), *The Benefits of Learning: the Impact of Education on Health, Family Life and Social Capital* (pp. 161–178). London: RoutledgeFalmer.
- Dohmen, G. (1998). Zur Zukunft der Weiterbildung in Europa: lebenslanges Lernen für alle in veränderten Lernumwelten. Bonn: BMBF.
- Ehlert, M. (2017). Who benefits from training courses in Germany? Monetary returns to non-formal further education on a segmented labour market. *European Sociological Review*, 33(3), 436–448.
- Einolf, C., & Chambré, S. M. (2011). Who volunteers? Constructing a hybrid theory. *International Journal of Nonprofit and Voluntary Sector Marketing*, 16(4), 298–310.
- Eisermann, M., Janik, F., & Kruppe, T. (2014). Weiterbildungsbeteiligung – Ursachen unterschiedlicher Teilnahmequoten in verschiedenen Datenquellen. *Zeitschrift für Erziehungswissenschaft*, 17(3), 473–495.
- Ekman, J., & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human affairs*, 22(3), 283–300.
- European Commission. (2002). *European report on quality indicators of lifelong learning*. Brussels: European Commission.
- Eurostat. (2017). Adult learning statistics – characteristics of education and training. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Adult_learning_statistics_-_characteristics_of_education_and_training.
- Feinstein, L., Hammond, C., Woods, L., Preston, J., & Bynner, J. (2003). *The contribution of adult learning to health and social capital*. Research Report No. 8. London: Centre for the Wider Benefits of Learning.
- Field, J. (2011). Researching the benefits of learning: the persuasive power of longitudinal studies. *London Review of Education*, 9(3), 283–292.
- Gesthuizen, M., Van der Meer, T., & Scheepers, P. (2008). Education and dimensions of social capital: do educational effects differ due to educational expansion and social security expenditure? *European Sociological Review*, 24(5), 617–632.
- Gorges, J. (2018). Weiterbildungsbeteiligung Älterer aus Perspektive der Erwartungswert-Theorie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 50(3), 149–159.
- Hanushek, E. A., Schwerdt, G., Woessmann, L., & Zhang, L. (2017). General education, vocational education, and labor-market outcomes over the lifecycle. *Journal of Human Resources*, 52(1), 48–87.

- Hustinx, L., Cnaan, R. A., & Handy, F. (2010). Navigating theories of volunteering: A hybrid map for a complex phenomenon. *Journal for the theory of social behavior*, 40(4), 410–434.
- Janoski, T., & Wilson, J. (1995). Pathways to voluntarism: Family socialization and status transmission models. *Social Forces*, 74(1), 271–292.
- Kaufmann-Kuchta, K., & Kuper, H. (2017). Informelles Lernen und soziale Teilhabe. In F. Bilger, F. Behringer, H. Kuper, & S. Schrader (Eds.), *Weiterbildungsverhalten in Deutschland 2016 – Ergebnisse des Adult Education Survey (AES)* (pp. 185–201). Bielefeld: wbv Media.
- Kleinert, C., Matthes, B., Antoni, M., Drasch, K., Ruland, M., & Trahms, A. (2011). ALWA–new life course data for Germany. *Schmollers Jahrbuch*, 131(4), 625–634.
- Kruppe, T., & Trepesch, M. (2017). Weiterbildungsbeteiligung in Deutschland. Auswertungen mit den Daten der Erwachsenenbefragung des Nationalen Bildungspanels „Bildung im Erwachsenenalter und lebenslanges Lernen“. *IAB-Discussion Paper*, 16/2017.
- Kuper, H., Christ, J., & Schrader, J. (2017). Formale Bildungsaktivitäten Erwachsener. In F. Bilger, F. Behringer, H. Kuper, & S. Schrader (Eds.), *Weiterbildungsverhalten in Deutschland 2016 – Ergebnisse des Adult Education Survey (AES)* (pp. 153–161). Bielefeld: wbv Media.
- Lancee, B., & Radl, J. (2014). Volunteering over the life course. *Social Forces*, 93(2), 833–862.
- Mania, E. (2019). Temporale Bezüge bei Regulativen der Weiterbildungsbeteiligung. *Zeitschrift für Weiterbildungsforschung*, 42(2), 201–219.
- Mascherini, M., Vidoni, D., & Manca, A. R. (2011). Exploring the determinants of civil participation in 14 European countries: One-size-fits none. *European Sociological Review*, 27(6), 790–807.
- Nie, N. H., Junn, J., & Stehlik-Barry, K. (1996). *Education and democratic citizenship in America*. Chicago: University of Chicago.
- Oesterle, S., Johnson, M. K., & Mortimer, J. T. (2004). Volunteerism during the transition to adulthood: A life course perspective. *Social Forces*, 82(3), 1123–1149.
- Preston, J., & Feinstein, L. (2004). *Adult education and attitude change*. Research Report No.11. London: Centre for the Wider Benefits of Learning.
- Priemer, J., Krimmer, H., & Labigne, A. (2017). *ZiviZ-Survey 2017 – Vielfalt verstehen. Zusammenhalt stärken*. Essen: SV Verwaltungsgesellschaft für Wissenschaftspflege mbH.
- Putnam, R. D. (2000). Bowling alone: America’s declining social capital. In L. Crothers & C. Lockhart (Eds.), *Culture and politics* (pp. 223–234). New York, NY: Palgrave Macmillan.
- Rotolo, T. (2000). A time to join, a time to quit: The influence of life cycle transitions on voluntary association membership. *Social Forces*, 78(3), 1133–1161.
- Rüber, I. E., & Bol, T. (2017). Informal learning and labour market returns. Evidence from German panel data. *European Sociological Review*, 33(6), 765–778.
- Rüber, I. E., Rees, S. L., & Schmidt-Hertha, B. (2018). Lifelong learning–lifelong returns? A new theoretical framework for the analysis of civic returns on adult learning. *International Review of Education*, 64(5), 543–562.
- Ruhose, J., Thomsen, S. L., & Weilage, I. (2019). The benefits of adult learning: Work-related training, social capital, and earnings. *Economics of Education Review*, 72, 166–186.
- Schnittker, J., & Behrman, J. R. (2012). Learning to do well or learning to do good? Estimating the effects of schooling on civic engagement, social cohesion, and labor market outcomes in the presence of endowments. *Social science research*, 41(2), 306–320.
- Schuller, T., & Desjardin, R. (2010). Wider benefits of adult education. In K. Rubenson (Ed.), *Adult Learning and Education* (pp. 294–298). Oxford: Elsevier.

- Schuller, T., Preston, J., Hammond, C., Brassett-Grundy, A., & Bynner, J. (2004). *The Benefits of Learning. The impact of education on health, family life and social capital*. London: Routledge Falmer.
- Simonson, J., & Gordo, L. R. (2017). Qualifizierung im freiwilligen Engagement. In J. Simonson, C. Vogel, & C. Tesch-Römer (Eds.), *Freiwilliges Engagement in Deutschland* (pp. 355–376). Wiesbaden: VS.
- Simonson, J., Hameister, N., & Vogel, C. (2017). Daten und Methoden des Deutschen Freiwilligensurveys. In J. Simonson, C. Vogel, & C. Tesch-Römer (Eds.), *Freiwilliges Engagement in Deutschland* (pp. 51–88). Wiesbaden: VS.
- Simonson, J., & Vogel, C. (2017). Organisationale Struktur des freiwilligen Engagements und Verbesserungsmöglichkeiten der Rahmenbedingungen. In J. Simonson, C. Vogel, & C. Tesch-Römer (Eds.), *Freiwilliges Engagement in Deutschland* (pp. 523–548). Wiesbaden: VS.
- Thomas, E. (2017). The outcomes and impacts of everyday learning. *International Journal of Lifelong Education*, 36(3), 308–323.
- Treiman, D. J. (2014). *Quantitative data analysis: Doing social research to test ideas*. Hoboken, New Jersey: Wiley.
- United Nations General Assembly. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. UN Doc. A/70/L/1. Retrieved from <https://sustainabledevelopment.un.org/post2015/transformingourworld>
- Van Ingen, E., & Dekker, P. (2011). Changes in the determinants of volunteering: Participation and time investment between 1975 and 2005 in the Netherlands. *Nonprofit and Voluntary Sector Quarterly*, 40(4), 682–702.
- Van Ingen, E., & van der Meer, T. (2011). Welfare state expenditure and inequalities in voluntary association participation. *Journal of European Social Policy*, 21(4), 302–322.
- Vera-Toscano, E., Rodrigues, M., & Costa, P. (2017). Beyond educational attainment: The importance of skills and lifelong learning for social outcomes. Evidence for Europe from PIAAC. *European Journal of Education*, 52(2), 217–231.
- Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University.
- Vogel, C., Hagen, C., Simonson, J., & Tesch-Römer, C. (2017). Freiwilliges Engagement und öffentliche gemeinschaftliche Aktivität. In J. Simonson, C. Vogel, & C. Tesch-Römer (Eds.), *Freiwilliges Engagement in Deutschland* (pp. 91–150). Wiesbaden: VS.
- Wiertz, D. (2016). Segregation in civic life: Ethnic sorting and mixing across voluntary associations. *American Sociological Review*, 81(4), 800–827.
- Wilson, J. (2012). Volunteerism research: A review essay. *Nonprofit and Voluntary Sector Quarterly*, 41(2), 176–212.

Appendix

Supplementary material – Robustness checks

Supplement A – Two wave linear regression

In the main analysis, I use cross-sectional data. It is, therewith, not fully possible to identify a causal relationship between adult learning and civil participation. One of the reasons is that in a cross-section the predictor does not necessarily precede the outcome. Since the NEPS is a panel study and adult learning behavior is part of the annual questionnaire, it is possible to include adult learning participation from the previous wave and thereby control for the time-order of predictor and outcome. I employ a two-wave multinomial logistic regression, wherein the predictor variables (adult learning) are constructed on the basis of the fourth NEPS-main wave, and where the outcome and controls stem from the fifth wave. The results are displayed in Table A. The estimates do not meaningfully change regarding the significance levels in comparison to the main model. But they do drop quite significantly in size. Nevertheless, the main implications hold.

Supplement B – Separate models per type of adult learning

In the main analysis, I simultaneously introduce formal, non-formal and informal learning to the model. This strategy might lead to skewed results since adult learners are likely to take part in more than one type of adult learning. Non-formal and formal learning participation may predict informal learning participation especially well. To rule out screwed results due to multicollinearity, I run the main models separately for the three different types of adult learning. The results are displayed in Table B. The coefficients of formal learning remain insignificant and are smaller (but positive) when compared to the main model. The coefficients of non-formal and informal learning are comparable to those in the main models, although they are slightly higher. The main implications hold.

Table A: Weighted two-wave multinomial logistic regressions of civil participation

| | Recurring civil participation | | Uptake of civil participation | |
|---|-------------------------------|---------------------|-------------------------------|---------------------|
| | 2 | 3 | 2 | 3 |
| Formal learning (<i>previous wave</i>) | 0.067 (0.19) | 0.328 (0.20) | 0.083 (0.20) | 0.233 (0.21) |
| Non-formal learning (<i>p.w.</i>) | 0.413*** (0.09) | 0.302** (0.10) | 0.262** (0.09) | 0.228* (0.09) |
| Informal: Reading (<i>p.w.</i>) | 0.587*** (0.10) | 0.574*** (0.10) | 0.185 (0.10) | 0.204 (0.11) |
| Informal: Media (<i>p.w.</i>) | -0.051 (0.11) | -0.011 (0.11) | 0.108 (0.12) | 0.136 (0.12) |
| Education (<i>Ref. = Low</i>) | | | | |
| Medium | 0.089 (0.13) | 0.195 (0.14) | -0.026 (0.12) | 0.023 (0.13) |
| High | 0.340** (0.13) | 0.387* (0.15) | -0.097 (0.14) | 0.019 (0.14) |
| Age | | 0.018*** (0.00) | | 0.013* (0.01) |
| Sex | | -0.159 (0.10) | | 0.015 (0.10) |
| Migration (<i>Ref. = Native</i>) | | | | |
| 1st generation | | -0.997*** (0.17) | | -0.252 (0.17) |
| 2nd generation | | -0.333* (0.14) | | -0.091 (0.15) |
| Child in HH | | 0.538*** (0.09) | | 0.422*** (0.10) |
| Health (<i>Ref. = Good</i>) | | | | |
| Medium | | 0.088 (0.11) | | -0.147 (0.11) |
| Poor | | -0.153 (0.20) | | -0.649** (0.20) |
| Household income | | 0.000*** (0.00) | | -0.000 (0.00) |
| Employment (<i>Ref. = Not employed</i>) | | | | |
| Part-time | | -0.156 (0.14) | | 0.034 (0.14) |
| Full-time | | -0.468** (0.13) | | -0.119 (0.15) |
| East | | -0.479*** (0.09) | | -0.069 (0.10) |
| Parents education (<i>Ref. = None</i>) | | | | |
| One | | 0.116 (0.12) | | -0.071 (0.143) |
| Both | | 0.215 (0.22) | | -0.552* (0.23) |
| Transitions | | | | |
| Divorced / Widowed | | 1.416 (0.77) | | 0.074 (0.58) |
| Birth | | -0.349 (0.27) | | 0.375 (0.31) |
| Married | | 0.128 (0.42) | | -0.587 (0.33) |
| Constant | -1.714*** (0.12) | -1.423*** (0.18) | -1.283*** (0.11) | -1.314*** (0.18) |
| <i>N</i> | 7870 | 7870 | 7870 | 7870 |

Note. Reference category is no civil participation. Multinomial logistic regression coefficients are displayed with standard errors in parentheses. Source: Own calculations using 4th and 5th orig. NEPS wave (SC6), version 8.0.0, doi:10.5157/NEPS:SC6:8.0.0.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table B: Weighted multinomial logistic regression models of civil participation, separate predictions per type of adult learning

| | Recurring civil participation | | | | Uptake of civil participation | | | |
|--|-------------------------------|---------------------|---------------------|---------------------|-------------------------------|---------------------|---------------------|---------------------|
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Formal | -0.036 (0.18) | | | | 0.097 (0.18) | | | |
| Non-formal | | 0.398*** (0.09) | | | | 0.411*** (0.08) | | |
| Informal: Reading | | | 0.821*** (0.09) | | | | 0.426*** (0.10) | |
| Informal: Media | | | | 0.404*** (0.10) | | | | 0.129 (0.10) |
| Education (<i>Ref.</i> = <i>Low</i>) | | | | | | | | |
| Medium | 0.329* (0.13) | 0.291* (0.13) | 0.222 (0.13) | 0.299* (0.13) | 0.095 (0.13) | 0.063 (0.13) | 0.039 (0.13) | 0.088 (0.13) |
| High | 0.678*** (0.14) | 0.624*** (0.14) | 0.442** (0.14) | 0.643*** (0.14) | 0.116 (0.14) | 0.070 (0.14) | -0.007 (0.14) | 0.109 (0.14) |
| Other controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Constant | -1.085*** (0.16) | -1.183*** (0.16) | -1.455*** (0.17) | -1.170*** (0.16) | -1.264*** (0.17) | -1.362*** (0.17) | -1.433*** (0.18) | -1.280*** (0.17) |
| <i>N</i> | 8317 | 8317 | 8317 | 8317 | 8317 | 8317 | 8317 | 8317 |

Note. Reference category is no civil participation. Multinomial logistic regression coefficients are displayed with standard errors in parentheses. Source: Own calculations using 5th orig. NEPS wave (SC6), version 8.0.0, doi:10.5157/NEPS:SC6:8.0.0.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Ann-Sophie Grub, Antje Biermann & Roland Brünken

Process-based measurement of professional vision of (prospective) teachers in the field of classroom management: A systematic review

Abstract

Effective classroom management is seen as a fundamental component of teachers' professional competence. The early detection of potential disturbances is of great importance for proactive control of the teaching process. Thus, professional vision serves as a link between teacher's knowledge and his or her actions in the event of deviations. Professional vision can be split into two aspects: noticing and reasoning. Previous research, based on subjective test procedures (i.e. video analysis or interviews), has primarily focused on the process of reasoning, whereas only a few studies have focused on the basal process of noticing, i.e. the recognition of possible disturbing situations. It is known from expertise research in different domains using process-based methods, such as eye-tracking, that experts and novices show differences in noticing processes. Therefore, to examine eye-tracking research for the teaching profession – especially noticing in classroom management – a systematic literature search was carried out between the years of 1999 to 2019. A total of 12 studies were found that recorded professional vision in the field of classroom management using eye-tracking. Overall, there were stable differences in the eye movement patterns of experts and novices for different parameters. However, some questions about the indicators used and possible influencing factors on expertise dependent perception remain unsettled.

Keywords

Professional competence; Professional vision; Noticing; Eye-tracking; Gaze behavior; Classroom management

Ann-Sophie Grub, M.Sc. (corresponding author) · PD Dr. Antje Biermann · Prof. Dr. Roland Brünken, Department of Education, Saarland University, Campus A4.2, 66123 Saarbrücken, Germany
e-mail: annsophie.grub@uni-saarland.de
a.biermann@mx.uni-saarland.de
r.brueken@mx.uni-saarland.de

Prozessbasierte Erfassung professioneller Wahrnehmung der Klassenführung bei (angehenden) Lehrkräften: Ein systematisches Review

Zusammenfassung

Effektive Klassenführung wird als grundlegender Bestandteil professioneller Kompetenz von Lehrkräften angesehen. Für eine proaktive Steuerung des Unterrichtsgeschehens ist das frühzeitige Erkennen von potentiellen Störungen von großer Bedeutung. Professionelle Wahrnehmung gilt als Bindeglied zwischen Wissen und Handeln der Lehrkraft und kann in die Aspekte Noticing und Reasoning unterteilt werden. Bisherige Arbeiten nutzten häufig subjektive Testverfahren (z.B. Interviews auf Basis von Videostimuli) zur Erfassung des Reasoning-Prozesses. Nur wenige Studien fokussieren auf den basaleren Prozess des Noticing. Aus der Expertiseforschung in unterschiedlichen Domänen, die prozessbasierte Methoden wie Eye-Tracking nutzen, ist bekannt, dass sich Novizen und Experten systematisch in der Erkennung potentieller Störsituationen unterscheiden. Das systematische Review gibt einen Überblick über die Arbeiten, die mit Eye-Tracking-Verfahren den Noticing-Prozess im Bereich der Klassenführung erfasst haben. Dafür wurde eine Literaturrecherche für den Zeitraum von 1999 bis 2019 durchgeführt. Insgesamt konnten 12 Studien identifiziert werden. Es zeigen sich stabile Unterschiede zwischen Experten und Novizen bei den meisten untersuchten Parametern. Sowohl die verwendeten Parameter als auch weitere mögliche Einflussfaktoren auf den Noticing-Prozess werden im Review diskutiert.

Schlagworte

Professionelle Kompetenz; Professionelle Wahrnehmung; Noticing; Eye-Tracking; Blickbewegungen; Klassenführung

1. Introduction

The professional competencies of teachers are the prerequisites for successful teaching, which is positively associated with student learning (Kunter et al., 2013). An important aspect of the multidimensional construct of competence is pedagogical knowledge, especially knowledge pertaining to classroom management (Kunter et al., 2013). This is viewed as a significant requirement for proactive control of the teaching process (Kounin, 2006) and therefore, is vitally important to the performance of students (Piwowar, Thiel, & Ophardt, 2013; Wang, Haertel, & Walberg, 1993). However, to be able to control teaching as proactively as possible, a clear professional vision of the teacher is necessary (cf. Blömeke, Gustafsson, & Shavelson, 2015). This ability describes how individuals observe and interpret events and situations in relation to their profession. Thus, professional vision can be divided into two aspects: noticing and reasoning (cf. Barth, 2017;

Seidel & Stürmer, 2014). The former is the necessary process of recognizing relevant cues, while the latter includes subsequent processes, such as interpreting what has been seen. Hence, the earlier a teacher perceives and anticipates situations relevant to classroom management, the better the proactive control of the teaching process will succeed, which can be summarized under the term “monitoring” (Gold, Hellermann, & Holodynski, 2017). Therefore, good knowledge organization is necessary (Wolff, 2015). Moreover, it is understood from expertise research that this knowledge base is more effectively organized (Wolff, 2015) among experts and thus, also influences perception (Stahnke, Schueler, & Roesken-Winter, 2016). As such, measuring methods for the early recognition of situations, especially for the reasoning process, video analysis, interviews, or questionnaires, have been widely used (cf. Seidel, Blomberg, & Stürmer, 2010). Within the field of expertise research, the eye-tracking method is already used in several domains (e.g. medicine, chess, aviation and traffic psychology) to capture cognitive perception processes (e.g. Gegenfurtner, Lehtinen, & Säljö, 2011), and can be harnessed to assess the noticing process in observing classroom management (e.g. Stürmer, Seidel, Müller, Häusler, & Cortina, 2017; van den Bogert, 2016). The use of this method has some advantages in comparison to verbal methods. At first, the recording of eye movements allows conclusions about the attention processes of the participants, which normally take place unconsciously and therefore, are difficult to verbalize (e.g. van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). Additionally, eye-tracking integrates both spatial and temporal information (Wolff, 2015), whereas with other methods, the focus can only be on one of these two tiers. To summarize, the advantage of eye-tracking can be seen in the direct and objective assessment of visual processing (cf. Wolff, 2015). It is also known from various domains, e.g. medicine, that these cognitive measures of behavior measurement are sensitive to differences in expertise (e.g. Gegenfurtner et al., 2011). Since eye-tracking methods offer the possibility to measure especially the basal ability of noticing (Riedl, Brandstätter, & Roithmayr, 2008), this paper presents an overview of different indicators of gaze behavior that can be used to capture differences in expertise. Moreover, the suitability of various indicators for measuring the professional vision of teachers in the field of classroom management is discussed.

2. Theoretical background

2.1 Professional competence

Teaching is a complex system characterized by multidimensionality, simultaneity, immediacy and unpredictability (Doyle, 1980). Accordingly, the high demands placed on teachers are also multifaceted. Therefore, research in the field of teaching emphasizes the importance of teacher competence; thus, the professional competence of teachers is a multidimensional construct (Blömeke et al.,

2015). Competence is assumed to be a continuum consisting of three components: the foundation is cognitive (e.g. professional knowledge, cf. Blömeke et al., 2015; Casale, Strauß, Hennemann, & König, 2016) and affective-motivational dispositions (e.g. aim of a broad activation of a learning group, cf. Blömeke et al., 2015; Casale et al., 2016). These dispositions influence the perception, as well as the interpretation of situations and the resulting decisions. Knowledge about classroom management as part of the generic dimension of pedagogical knowledge affects the ability to perceive and anticipate potential disruptive behavior and thus, the organization of disturbance-free lessons (Wolff, 2015). Especially for novice teachers, it is often a challenge to perceive all critical moments and to show effective handling of imminent or occurring disorders due to multidimensionality, as well as the simultaneity of interactions and events in the classroom and the need for immediacy of reactions (Wolff, 2015).

2.2 Classroom management

The basis for effective classroom management is the smooth organization of all activities in the classroom and to avoid interruptions and misconduct to maximize the time for teaching and learning (Emmer & Stough, 2001; Kounin, 2006; Steffensky, Gold, Holodynski, & Moeller, 2015). Therefore, visual monitoring and scanning of the classroom are of high importance (Gold & Holodynski, 2017). With the terms “withitness” and “overlapping,” Kounin (2006) identified two skills that are particularly crucial for professional perception. Withitness describes the ability of a teacher to be informed at all times about what is happening in the classroom (cf. multidimensionality, Doyle, 1985). To manage the classroom adequately and effectively, a precise detection and observation of relevant cues and events is required (van den Bogert, van Bruggen, Kostons, & Jochems, 2014; Wolff, 2015). Overlapping means that the teacher is able to deal with different issues at the same time, e.g. flow of instruction and some unavoidable disciplinary problems (cf. simultaneity; Doyle, 1985) to react appropriately to the situation (cf. immediacy; Doyle, 1985). Both abilities are dependent on constant visual attention processes (van den Bogert et al., 2014). The dimensions of withitness and overlapping can be summarized under the term “monitoring” (Gold & Holodynski, 2017). Thus, monitoring is a process that encompasses the teachers’ awareness of all relevant and simultaneously operating processes in the classroom, as well as the demonstration of such awareness to students (Gold & Holodynski, 2017). Therefore, perception, especially monitoring, should be a key component of any teacher because early detection of any student disturbances and the awareness of what is happening in the classroom are relevant factors for classroom management (van den Bogert et al., 2014).

2.3 Professional vision

There are several models to explain perception within the professional context (Endsley, 1995; Goodwin, 1994; Seidel et al., 2010; Seidel & Stürmer, 2014; Sherin, 2001). Two modeling approaches (situation awareness; Endsley, 1995, and professional vision; Sherin, 2001) focus on the perception process in a professional context, especially in teaching. Both models of professional vision describe the process of perception with sub-processes. Endsley (1995), in the model of *situation awareness*, describes three sub-processes (“perception,” “comprehension,” and “anticipation” to further events). This model originated from the domain of dynamic human decision-making systems and can be applied to the domain of teaching (Barth, 2017). The three levels are built upon each other and it is assumed that the level of difficulty in mastering each level increases from level 1 to level 3. As such, the perception of relevant situations is influenced through individual (e.g. experience, ability) and environmental factors (e.g. complexity, stress). *Professional vision* is a concept derived from Goodwin (1994) and adapted by Sherin (2001). The process of professional vision can also be divided into three sub-processes (“identification,” “making connections,” and “reasoning”), which many authors summarize into two main processes, including “noticing” (“identification”) and “knowledge-based reasoning” (summarizing “making connections” and “reasoning”) (e.g. Barth, 2017; Seidel et al., 2010; Seidel & Stürmer, 2014). Noticing describes the ability to focus attention on situations that are relevant for teaching and learning, whereas knowledge-based reasoning describes the ability to apply knowledge about teaching and learning and to draw appropriate conclusions. This process can be seen as an indicator of the quality of the application of knowledge to the situation in the classroom. Noticing and knowledge-based reasoning interact with each other (Sherin, 2007). Another similarity of the two models is that the process of identifying relevant cues or events is seen as the necessary process of professional vision. Without this inevitable process of noticing, processes of reasoning based on it cannot take place. In the following, the term “professional vision” will be used throughout, as this is the more common term in education research (c.f. Barth, 2017; Gold & Holodynski, 2017).

The competence model of Blömeke et al. (2015) also ascribes a unique role to perception because this process represents an essential basis for situation-specific skills. The interpretation of situations and the relevant cues and decision-making is possible only on the basis of perception. Therefore, noticing represents an unavoidable step in the perception of aspects of teaching relevant to classroom management.

2.4 Expertise

Expertise differences can be found in various domains, such as aviation, chess, and medicine, as well as in education (Gegenfurtner et al., 2011; Gegenfurtner & Seppänen, 2013; Reingold & Sheridan, 2011; Stahnke et al., 2016; van der Gijp et al., 2017). For example, Housner & Griffey (1985) found expertise differences in terms of classroom perception. According to Berliner (2001), expertise is domain specific. That is, knowledge is better structured by experts in the specific domain than by novices. Problems are represented in qualitatively different ways than by novices so that experts can remember meaningful schemata faster, e.g. experienced teachers are expected to process information faster and thus, need less time to comprehend a classroom situation (Chi & Glaser, 1988). Besides, experts are flexible and opportunistic planners who can change representations faster. In contrast, novices are more rigid in their concepts. Furthermore, experts are able to interpret ambiguous stimuli, while novices are more easily deceived. Thus, it can be summed up that experts tend to steer “top-down,” while novices tend to be subject to “bottom-up” cognitive processes (Hershler & Hochstein, 2009). Also, Livingston and Borko (1989) claim that the distinction between expert and novice teachers occurs in the flexibility of cognitive schemata. Similarly, the model by Wolff (2015) describes the influence of knowledge (“classroom management scripts”) on expert and novice teachers’ perceptions, awareness and interpretative processing of problematic classroom situations. It has been found that teachers’ knowledge organization affects the perceptions of classroom situations and ensuing situational awareness.

To distinguish experienced teachers from novice teachers, the criteria of Palmer, Stough, Burdinski, and Gonzales (2005) can be used. Palmer et al. (2005) postulated that experts should be selected by years of experience (a minimum of three to five years), social recognition (two or more different socially recognized persons from the corresponding domain), professional or social group membership, and performance-based criteria (e.g. student achievement).

To measure differences in expertise in the visual perception of situations relevant for classroom management, process-based methods are particularly suitable. This unique appropriateness results from the fact that visual perception is a continuous process. Over recent years, eye-tracking has become an important tool to investigate such processes (Holmqvist et al., 2011) because most information is processed through the eyes and these sensory organs are therefore ascribed a special role in the process of visual perception (in particular, information reception and processing) (Jarodzka, Holmqvist, & Gruber, 2017).

2.5 Acquisition of professional vision

Research on teacher expertise often uses high inference or subjective measurement methods because findings often rely on the interpretation of, for example, traditional observational data, video analysis, or interviews (Cortina, Miller, McKenzie, & Epstein, 2015). In a study from Wolff, van den Bogert, Jarodzka, and Boshuizen (2015), participants were encouraged to describe short video vignettes regarding classroom management, while Carter, Cushing, Sabers, Stein, and Berliner (1988) used photographic slides from classrooms, urging participants to describe and provide a reconstruction of the situation. In a study of Huang and Li (2012), the participants were asked to describe two lessons and then make suggestions for improvement and an evaluation of the watched lessons.

These studies could uncover differences between experts and novices in their perception and interpretation of situations relevant for classroom management on the reasoning level; however, these studies could not make a prediction about exactly which visual attention processes are responsible for these differences. Expertise research of different domains (medicine, chess, etc.), on the other hand, are traditionally based on low-inference objective measures (such as the measure of gaze movements) as an indicator of cognitive functions (van Gog et al., 2009). Eye-tracking has long been used to investigate selection and attention patterns (Bucher & Schumacher, 2012) and to obtain online measures of cognitive processes (Lachner, Jarodzka, & Nückles, 2016). This allows gaze movements to be used as behavioral indicators for cognitive processes (e.g. eye fixation data reflect attention and shifts in attention, cf. van Gog et al., 2009). As such, eye-tracking offers the possibility to investigate teachers' basic perceptual processes on the noticing level (Wolff, Jarodzka, van den Bogert, & Boshuizen, 2016).

2.6 Parameter/Indicators of professional vision

To assess professional vision with the eye-tracking method, several indicators are worth considering. According to the oculomotor definition, a fixation describes the period of time in which the eye behaves relatively motionless, i.e. moves as little as possible (Holmqvist et al., 2011). In general, fixations are an indicator, such as which environmental areas are allocated with attention and from which areas information is received or rather which stimuli are important (Just & Carpenter, 1976). Thus, the degree of experience influences the *number of fixations* (also called fixation density). Experts have more fixations on relevant areas in their specific domain, such as aviation (Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001), as well as dynamic stimuli (Jarodzka, Scheiter, Gerjets, & van Gog, 2010). This indicates that experts use their elaborated schemata while perceiving scenes, including school lessons. Teaching experts have theories about how to recognize potential disorders early and fixate exactly these areas more often (e.g. students who are not attentive or play with extra-curricular teaching materials). *Fixation dura-*

tion describes the period of time that a fixation continues and is likely to be the most used measure in eye-tracking research (Holmqvist et al., 2011). Jarodzka et al. (2010) demonstrated that novices have longer mean viewing times than experts. This implies that experts are faster in encoding information due to their professional knowledge (Chi & Glaser, 1988), which allows individuals with higher expertise to plan ahead in dynamic situations (cf. Jarodzka et al., 2017). Based on these eye movement parameters (many and short fixations), dynamic classroom situations can be updated permanently (“Like circus performers who keep plates spinning on top of sticks, teachers must not only establish a management system that works but keep it working by monitoring events continually and responding when breakdowns occur.”, Brophy, 1988, p. 3). Another measure is the *distribution of fixations*. In general, the distribution of fixations can be an indicator for deeper cognitive processing or the importance of a region (Kuperman, Bertram, & Baayen, 2008; Reingold, Charness, Pomplun, & Stampe, 2001), which leads to a relatively longer fixation time for relevant regions for experts in a domain. Regarding monitoring behavior or the awareness of what is going on in the classroom, the distribution of attention across students should be more evenly placed for expert teachers (Wolff, 2015).

In the context of monitoring, two further parameters can be used: *revisits* and *fixation skips*. Repeated fixations respectively revisits give insight about gathered and updated information (cf. Wolff, 2015). Through the analysis of areas in which people look back, conclusions can be drawn about which areas are classified as relevant by the person, whereas fixation skips highlighted areas, which are ignored from viewers. This can also be used to identify which areas, for example for experts, are classified as rather irrelevant due to their knowledge and are therefore visually skipped, and which areas, in contrast, are particularly relevant and are therefore reviewed permanently again and again.

When interpreting eye movements as a method for recording underlying cognitive processes, some assumptions are often made implicitly or explicitly (Rötting, 2001). According to Just and Carpenter (1980), for example, it is assumed that there is no appreciable lag between what is fixed and what is processed (eye-mind-assumption). This variant of interpretation has been criticized for a long time (e.g. Ehrlich & Rayner, 1983). The interpretation of the parameters and the use of the parameters as indicators should therefore in principle be done on a theoretical basis and never on the basis of a measure alone (cf. Murray, Fischer, & Tatler, 2013).

The aim of the study is to provide an overview of the current research on the process-based indicators used thus far to record professional vision among experts and novices in the field of classroom management.

3. Method

A systematic review (van Wee & Banister, 2016) of the literature from 1999 until 2019 was performed by using three databases, including ERIC, PsycINFO and the Web of Science. At first, the databases were searched by using different combinations of relevant keywords included in the full text of the article. Those keywords included expertise, teach, classroom management and eye-tracking, as well as related words and synonyms (a comprehensive list of search strings is attached in the Appendix, Table A). To minimize publication bias, not only different databases were used, but search restrictions (e.g. peer review procedures) were also dispensed with (cf. Zawacki-Richter, Kerres, Bedenlier, Bond & Buntis, 2020). This search revealed a sample of 1,178 papers after the elimination of duplicates. To be included in the database, during a second step, the studies needed to contain the examination of the perception of classroom management from (pre-service) teachers through process-based measurements. Therefore, all papers were excluded whose title and abstract content did not coincide with the research topic. This resulted in 29 articles, which were examined more closely by two reviewers during a third step. Thereafter, all papers were excluded whose content did not coincide with the research topic and in particular, the process-based acquisition of perception, resulting in a total of eight papers. In the last step, all remaining papers were cross-referenced, which led to three additional articles. In total, the database search and cross-references identified 11 articles with 12 studies that were included in the review. Thereof, three studies came from doctoral theses, which have not yet been published in peer-review journals (a graphical representation of the complete process can be found in Appendix, Figure A). For a more detailed analysis, information on the sample, the selection criteria of experts and novices, stimuli, as well as the eye-tracking method used were extracted in Table 1. Moreover, the parameters used to assess professional vision and their indication in classroom management were extracted to analyze them more precisely in the results chapter (see Table 2).

4. Results

4.1 Methods and equipment

We found eight studies using a real world classroom with mobile eye-tracking (1, 2, 3, 4, 5, 6, 11, and 12; see Table 1), and four studies using videos from lessons as study materials with a static eye-tracker (7, 8, 9, and 10; see Table 1). Eye-tracking records in real classrooms have durations between 10 and 45 minutes. However, eye-tracking records in a laboratory setting were based on different numbers of videos (1 to 8) and different duration times (about 2 to 4 minutes).

Table 1: Descriptive description

| | Author (year) | Method | Stimulus | Sample (N) | | | Expertise criteria | |
|---|--------------------------------------|---------------------|----------------------------------|------------|-----|--|--|--|
| | | | | E | N | E | N | |
| 1 | Cortina et al. (2015) | Mobile eye-tracking | Real classroom | 12 | 12 | - Mentors | - Mentees | |
| 2 | Dessus, Cosnefroy, & Luengo (2016) | Mobile eye-tracking | Real classroom | 2 | 2 | - 20 and 25 years teaching experience | - 0.5 years teaching experience | |
| 3 | McIntyre & Foulsham (2018) | Mobile eye-tracking | Real classroom | 20 | 20 | - At least 6 years' experience, - Social nomination, - Professional membership, - Performance rating | - Teachers in the same school who con- - trusted most with the expertise criteria | |
| 4 | McIntyre, Jarodzka, & Klassen (2017) | Mobile eye-tracking | Real classroom | 20 | 20 | - cf. McIntyre & Foulsham (2018) | - cf. McIntyre & Foulsham (2018) | |
| 5 | McIntyre, Mainhard, & Klassen (2017) | Mobile eye-tracking | Real classroom | 20 | 20 | - cf. McIntyre & Foulsham (2018) | - cf. McIntyre & Foulsham (2018) | |
| 6 | Stürmer et al. (2017) | Mobile eye-tracking | M-Teach situations | / | 7 | - / | - Preservice teachers | |
| | | | Real classroom | 4 | - / | - / | - / | |
| 7 | van den Bogert et al. (2014) | Eye-tracking | Videos from real life classrooms | 7 | 7 | - Selected by three teacher-supervisors - 10 years of experience teaching, - Possess a keen understanding of what goes on in their classrooms, - Are known to create and maintain a positive learning climate | - Randomly selected from first and second year student teachers | |

Table 1 continued

Table 1 continued

| Author (year) | Method | Stimulus | Sample (N) | | Expertise criteria |
|---|---------------------|----------------------------------|------------|----|---|
| | | | E | N | |
| 8 van den Bogert (2016) | Eye-tracking | Videos from real life classrooms | 34 | 32 | – cf. van den Bogert et al. (2014) – First and second year student teachers, – 10-40 hours of teaching experience |
| 9 Wolff, Jarodzka, van den Bogert, & Boshuizen (2016) | Eye-tracking | Videos from real life classrooms | 35 | 32 | – 7 years teaching experience, – Recognized by fellow teachers as a competent or above-average classroom manager, – Recommended by their school leaders as experts in teaching profession |
| 10 Yamamoto & Imai-Matsumura (2013) | Eye-tracking | Video from real life classroom | 43 | / | – From in-service teacher training, – 15.7 years, SD = 8.7 teaching experience |
| 11 Huang (2018) | Mobile eye-tracking | Real classroom | 25 | 25 | – Experienced teachers – Mentees |
| 12 Huang (2018) | Mobile eye-tracking | Real classroom | 10 | / | – / |

Note. E = experts, N = novices.

4.2 Sample

Eight of the studies compared the gaze behavior between expert and novice teachers (1, 2, 3, 4, 5, 7, 8, 9, and 11; see Table 1). Concerning the selection criteria of the experts and novices, different levels of detail were given in the studies. Studies 1, 11, and 12 described experts simply as mentors or experienced teachers and the novices as mentees but did not provide any information on teaching experience. Contrary, the other studies (2, 3, 4, 5, 7, 8, and 9) used the length of teaching experience as a criterion for distinguishing between experts and novices; thus, experts had a minimum of six years of experience. Novices, on the other hand, had only 0.5 years or 10-40 hours of teaching experience. Multiple studies used further criteria to distinguish between experts and novices (3, 4, 5, 7, and 9), such as social nomination or classroom management performance. In Study 10, the length of teaching experience was used as a continuous variable, but no novices were included. In Study 6, only novice teachers were included, while in Study 12, only expert teachers attended, teaching different subjects. Both studies gave critical insights into cognitive structure and possible influencing factors beyond the expertise of teachers' gaze and therefore, were included in the review.

4.3 Parameter and results

4.3.1 Number of fixations

Five of the 12 analyzed studies used the parameter “number of fixation,” while Study 7 revealed more overall fixations for expert teachers (see Table 2). Studies 8 and 11 focused on relevant (e.g. disorderly behavior of students, instructional material) vs. irrelevant areas, and both studies showed that both groups had a comparable total number of fixations, such as in Study 7. However, the results from Studies 7, 8 and 11 also showed that experts have more fixations these relevant areas more often than novices. Results from Study 10 (only expert teachers) displayed that teachers fixated on the disturbing students more often when they were aware of them than when they were not. Unexpectedly, the aware teachers did not have longer teaching experience than the unaware teachers, which eventually was explained with a problematic operationalization of disturbing behavior in the study (see also Wolff et al., 2016). Study 6 (only novices) demonstrated that the participants mostly looked at the students, followed by the instructional material.

In conclusion, the three studies with the expert-novice-comparison, indicate that experts have more fixations than novices, especially on relevant areas or rather, students in the classroom, which give first insights, that the assumptions can be confirmed with the eye-tracking method, too. Further, the results of Study 6 indicate that novices focus mostly on students and may be aware of the relative importance of students as cues for effective classroom management.

4.3.2 Fixation duration

Eight studies used the parameter “fixation duration” (see Table 2). Results from Study 7 and 11 revealed that experts have overall shorter fixation durations compared to novices; however, in Study 8, only with a more fine-grained expertise-classification with four expert groups could differences be seen between an inverse U-shaped relationship for fixation duration. Thus, conclusions from these studies indicated, that by the use of eye-tracking, results can be achieved which correspond to the assumptions, such as fixation duration can be seen as an indicator of the fast encoding processes by experts (Chi & Glaser, 1988). Results from Study 5 indicated that experts have longer durations directed at students compared with novices. Contrary, novices had longer durations on non-student areas than experts. The outcomes of Study 2 using gazing time showed that every teacher looked at every student a few times but experts in this study focused longer on less abled students (in terms of behavior) than novices. Study 8 found no differences in the total dwell time regarding disruptive students between the expert groups (focus on length of dwell time). A possible explanation for this is the high salience of these pupils in the selected scenes. Participants looked longer at the students, followed by the instructional material; nevertheless, in Study 10, the teachers who were aware of the disturbing students watched the troublemakers longer than the other students. These results are also in line with the assumption of deeper processing by focusing on essential areas (Kuperman et al., 2008; Reingold et al., 2001). When looking at different subjects taught by the same teacher (Study 12), the observed fixation duration on students was longer when teaching literacy compared to math. This implies that this indicator not only depends on expertise but also on the issues and goals of the specific subject.

4.3.3 Fixation dispersion

Ten studies used this relative parameter with different measures (see Table 2): Gini-coefficient (1 & 2), gaze proportion (4 & 5), ranked scores of total fixation frequency and fixation duration (6 & 7), total dwell time (8, focus on dispersion of dwell time), fixation dispersion average (9), standard distance deviation (11) and Gaussian density plot (12). Studies 1, 2, 6, and 7 investigated distribution across different students in a classroom as an aspect for monitoring behavior. Studies 4, 5, 8, 9, 11 and 12 were more interested in the distribution of the gaze behavior across more didactical aspects in the classroom (e.g. learners in relation to instructional material) as an indicator for their relevance regarding instruction, especially classroom management. In general, results from the first mentioned studies indicate that experts scatter their view more evenly across students; that is, that they show better monitoring behavior. However, two of these studies resulted in a more differentiated look. Study 2 also assessed information from students regarding their behavior and their academic achievement. At a descriptive level, the au-

thors could show that experts searched for more information regarding students with problematic behaviors in relation to other students, while novices distributed their attention more evenly across students. In Study 6, the novices taught a class with four students with defined roles (cf. struggling, uninterested). Overall, for novice teachers, the distribution over the four students was relatively even but the results also show strong interindividual differences between teachers.

Regarding distribution over didactical events, the results of Study 4 show that experts look more frequently at relevant aspects (e.g. students) in comparison to other didactical aspects (like instructional material) or irrelevant stimuli, while novices distributed their attention more evenly across non-instructional material and students (focus on distribution of gaze behavior). Study 9 also showed that experts were more focused when taking an overall distribution of fixations into account. In contrast, in Study 11, the dispersion was higher for experts. Unfortunately, the authors of both studies gave no deeper insights about the precise content of the focus. However, Study 12 further takes into account that interpretation of the fixation dispersion not only depends on the analytical unit (cf. distribution over students or over didactical aspects) but also on different instructional aims in different subjects. In regard to literacy, the focus remained more on students, whereas in math, the focus was set more on instructional material, such as the blackboard, relative to the other elements.

4.3.4 Skips and revisits

Only Study 9 examined the skips in eye-movement behavior (see Table 2). Descriptive results suggested that novices skipped more areas than experts, though the authors found no statistical interference confirmation. Studies 5 and 9 engaged with revisits as an indicator for monitoring, while results from Study 9 demonstrated that experts revisited more areas than novices when observing classroom scenes. Thus, it appears that experts searched for activity between students, as well as following posture and body movements. Results from Study 5 supported these findings, along with shorter return times to relevant attractors for experts.

4.3.5 Others

Some studies used further parameters to examine eye-movement behavior (see Table 2). For example, Study 3 applied a string based scanpath analysis, which preserves information about what and in which order the participants looked, as well as what could be seen as a reflection about the internal representations of the teachers. Due to a shared structure of professional knowledge, the internal model and scanpaths of experts should be equal compared with novices. In line with this assumption, the gaze behavior of different teachers at one level of expertise is more similar than the gaze behavior of different teachers at various levels of expertise.

Table 2: Methodological description

| | Author (year) | Parameter | Indicator for | Hypothesis | Confirmation |
|---|--------------------------------------|---|--|-----------------------|------------------------------|
| 1 | Cortina et al. (2015) | GINI-coefficient | Monitoring | $E < N$ | Y |
| 2 | Dessus, Cosnefroy, & Luengo (2016) | GINI-coefficient | Monitoring | $E > N$ | Y |
| | | Gazing time | Attention | $E > N$ | Y |
| 3 | McIntyre & Foulsham (2018) | Scanpath | Monitoring | $\Delta E < \Delta N$ | Y |
| 4 | McIntyre, Jarodzka, & Klassen (2017) | Gaze proportion | Attention | $E > N$ | Y |
| 5 | McIntyre, Mainhard, & Klassen (2017) | Gaze duration | Cognitive processes | $E > N$ | Y |
| | | Return times | Efficiency | $E > N$ | Y |
| | | Gaze flexibility | Monitoring | $E > N$ | Y |
| 6 | Stürmer et al. (2017) (only N) | Number of fixation | Attention | High σ | Y |
| | | Fixation duration | Attention | High σ | Y |
| | | Switchover | Attention | High σ | Y |
| | | Ranked scores of total fixation frequency | Monitoring | High σ | Y |
| 7 | van den Bogert et al. (2014) | Mean fixation duration | Processing time | $E < N$ | Y |
| | | Mean fixation count | Monitoring | $E > N$ | Y |
| | | Total time fixation duration scores | Monitoring | $\Delta E < \Delta N$ | Y |
| 8 | van den Bogert (2016) | Mean fixation duration | Processing time | $E < N$ | N (2 groups) Y (4 groups) |
| | | Total dwell time | Divide attention between one salient event and the rest of the classroom | $E < N$ | N ($E \neq N$) |
| | | Scanpath length | Monitoring | $E > N$ | Y |

Table 2 continued

Table 2 continued

| Author (year) | Parameter | Indicator for | Hypothesis | Confirmation |
|---------------|---|-------------------------------|------------|------------------------------|
| 9 | Wolff, Jarodzka, van den Bogert, & Boshuizen (2016) | Monitoring | $E > N$ | Y |
| | Number of visits | Monitoring | $E \neq N$ | Y (E > N) |
| | Fixation dispersion average | Monitoring | $E \neq N$ | Y (E > N) |
| | Number of revisits | Gather and update information | $E \neq N$ | Relevant areas: E > N |
| 10 | Yamamoto & Imai-Matsumura (2013) (only E) | Ignored classroom information | $E \neq N$ | N (E = N) |
| | Fixation count | Awareness | A ≠ UA | Y (A > UA) |
| | Fixation length | Awareness | A ≠ UA | Y (A > UA) |
| | Mean fixation duration | Awareness | A ≠ UA | N (A ≅ UA) |
| 11 | Huang (2018) Study 1 | Time to first fixation | A ≠ UA | N (A ≅ UA) |
| | Distribution of fixations | Attention | E > N | Y |
| | Fixation duration | Attention | E < N | Y |
| | Fixation count | Monitoring | E > N | Y |
| 12 | Huang (2018) Study 2 (only E) | Saccade amplitude | / | E > N |
| | Saccade direction | Monitoring | / | E ≠ N |
| | Distribution of fixations | Attention | / | Math = literacy |
| | Fixation duration | Monitoring | / | On students: literacy > math |
| | Fixation count | Monitoring | / | On students: literacy > math |

Note. E = experts, N = novices, A = aware teachers, UA = unaware teachers.

Study 8 used scanpath length, which is defined according to Holmqvist et al. (2011) as the sum of all saccadic amplitudes in a scanpath. The study found a positive linear relationship between the level of expertise and scanpath length, which can be interpreted as a broader monitoring behavior for experts. In Study 5, the authors used a state space grid-method to examine gaze behavior on a more didactical level. With this method, the authors could define a so-called “efficient didactical gaze” (the most stable and prevalent gaze behavior) and calculated the return time to this gaze for experts and novices. Results showed that experts returned faster to the efficient gaze. Moreover, the authors measured the flexibility of gaze behavior regarding various instructional aims and showed that experts adapted their gaze behavior. That is, experts had different amounts of transitions between student and non-student areas depending on their aims (transmit information to students vs. gain information about students).

5. Discussion

This review aimed to systematize the current research in the field of professional vision among experts and/or novices regarding classroom management measured with eye-tracking. Based on a literature search, 12 studies have been included and described concerning the used indicators, their sensitivity to expertise differences, and further influencing factors.

5.1 Methods

The included studies differ regarding the applied stimuli and their standardization. Several studies presented videos of school lessons relying on high controllability and standardization, while other studies preferred higher external validity through the use of mobile eye-tracking in real classrooms. Results from Foulsham, Walker, and Kingstone (2011) suggest that eye-tracking data from the real world in comparison to data from videos differ at some points. For example, the condition of the real world allows a head turn and thus, the natural freedom of movement. Furthermore, real classrooms differ in many aspects, which makes comparisons more complicated, including the composition of students, the arrangement of the material elements, and the general learning situation. Some of the studies with real classrooms increased the standardization at some points; for instance, Study 6 used a highly standardized simulated situation with the same students being given roles. In Study 11, the respective pairs (mentors & mentees) taught the same students with the same topic in the same classroom, while in Study 12, the same teacher taught two subjects in the same class. Nevertheless, systematic differences in the expertise of eye-movement behavior could be uncovered both in a real world classroom and in a laboratory setting.

Furthermore, the use of various eye-tracking devices and different technical features can present an influencing variable on the sensitivity of the measurements (in the analyzed studies e.g. different sampling rates with a range from 30 Hz to 250 Hz were used). However, differences in expertise in gaze behavior were found across all studies, which suggests that this behavior differs between experts and novices and that the eye-tracking procedure is sensitive to these differences.

5.2 Sample characteristics

Several of the studies with an expert-novice-paradigm distinguished experts and novices only on the basis of years of experience, while other studies applied at least one other criterion postulated by Palmer et al. (2005). These factors included social recognition, professional or social group membership, though only for experts. This can be problematic in several ways, as some novice teachers can have a well-developed professional vision (cf. due to former experience with students or children, see Studies 7 and 8). Moreover, those deemed expert teachers based only on their years of experience may not necessarily be good perceivers (see Study 10; also Palmer et al., 2005). Thus, studies should apply the criteria by Palmer et al. (2005) and control professional knowledge for the investigated aspects of professional vision (cf. Lachner et al., 2016).

Furthermore, it should be kept in mind that the sample size of some of the included studies were relatively small with limited power for inferential statistics.

5.3 Differences in parameter used and their sensitivity to expertise

Two of the parameters most commonly used in the papers were “number of fixations” and “fixation duration.” It was found that generally, experts have more, but shorter fixations, whereas novices have less, but longer fixations. This is in line with the assumption of fast encoding processes by various experts (Chi & Glaser, 1988). Specifically, experts fixated relevant areas (e.g. disruptive students) more often and for a longer duration than irrelevant areas (e.g. non-instructional material) compared to novices, who looked more frequently at irrelevant areas. These results confirm deeper cognitive processes by focusing on important areas (Kuperman et al., 2008; Reingold et al., 2001). However, mere eye movement data should always be interpreted with caution (see the section of further implications).

In almost all included studies, parameters were used to measure attention distribution but with different indicators (e.g. Gini-coefficient, difference coefficients, fixation duration, and fixation frequency). These parameters are useful to assess monitoring behavior or to assess the relevance of one student in a group of students (e.g. a disruptive student) or didactical objects regarding the teaching situation (e.g. learners vs. instructional material). It was found that experts distributed

their attention more evenly over students and spent, on average, less time with single area of interest (AOI). This result can be combined with the results of the number of fixations and fixation duration, as fast and short fixations allow permanent and fast scanning of the classroom. This ability to monitor has also been recorded e.g. utilizing previously fixed AOIs. This reveals that experts return more frequently to relevant areas and therefore, can regularly update their internal information about what is happening throughout the classroom. Furthermore, scanpath comparisons indicate that more experienced teachers look around more frequently in the classroom, i.e. experts scan the classroom more actively than novices. This result also matches the expertise-dependent observation behavior of monitoring (cf. Brophy, 1988; Carter et al., 1988).

Using dispersion measures often entails the problem of inappropriate aggregation of data (Orquin & Holmqvist, 2018). A stronger focus on one area in relation to another may first be due to higher relevance recognized by top-down processes (cf. Orquin & Mueller Loose, 2013), second due to the complexity of the object (Just & Carpenter, 1976), or third due to its saliency (bottom-up processes, Itti & Koch, 2001). To this end, it is important that a description of the meaning (e.g. students, student material) and of the relevance of the AIOs for the teaching process is of great importance for interpreting the data. This process should not only be made data-driven (a relevant AOI is one that experts focus on more often and for longer periods of time), but also normatively by the researcher or both regarding a special teaching situation (see also the next section on implications for further research).

However, there are also limitations regarding the expertise-sensitivity of the eye-tracking parameters in the context of professional vision. Five of the included studies (3, 4, 5, 10, and 12) concluded that expertise is not the only influence on teachers' visual processing but also their cultural background or the subject. Thus, it is known that the cognitive models of teachers depend on shared culture (Blömeke, Olsen, & Suhl, 2016; Hofstede, 1986). However, didactical strategies and the teachers' gaze behavior or the relevance of stimuli depend on the aims of the subject (König, Blömeke, Paine, Schmidt, & Hsieh, 2011). According to the Endsley's model of situation awareness (1995), besides expertise, for example, stress (e.g. Sneddon, Mearns, & Flin, 2013), cognitive load (e.g. Dessus, Cosnefroy, & Luengo, 2016; Prieto, Wen, Caballero, Sharma, & Dillenbourg, 2014) or task characteristics (e.g. Gegenfurtner et al., 2011; Yarbus, 1967) represent influencing variables during the perception process.

5.4 Implications for further research

Over the past decade, eye-tracking has become increasingly popular as a method for educational research. In turn, the equipment needed to study this method has become cheaper, while the handling has become easier (Orquin & Holmqvist, 2018). As a consequence, many test subjects can be collected with relatively little effort, making it possible to develop more complex designs with several comparison groups that still have sufficient power.

One question on the definition of expertise is at what point professional vision starts to develop and if the development of professional vision is similar to the five-stage model of expertise acquisition postulated by Dreyfus (2004). Thus, the use of these intermediate stages of expertise development in relation to professional vision could possibly explain the inverse U-shaped connection found by van den Bogert (2016).

Potentially, the use of standardized videos or scenarios would be a suitable method for future studies to experimentally manipulate the extent of the complexity caused by classroom management events. It could be investigated whether the differences in expertise are expressed differently depending on the number and integrity of events in the classroom (cf. Wolff et al., 2016). Contrasting videos with situations relevant for classroom management without such situations could also provide information on whether, and if so, the gaze behavior of experts and novices relates to classrooms without disturbing elements.

With regard to the method of eye-tracking, there are also limitations, with regard to the evaluation and interpretation of the parameters. The eye-mind-assumption (Just & Carpenter, 1980) states that eye fixations are directly linked to attention processes. Shifts in fixation are therefore also directly linked to shifts in attention (Holmqvist et al., 2011). It is generally assumed that a person who fixates on a particular point also processes information contained in that point simultaneously. However, this assumption does not take into account, for example, hidden attention (Posner, 1980): It is possible to look at a point, i.e. to fixate on it, without actually perceiving it. It is also possible to stare emptily into space and think about completely different things than about the point that is fixed. Covert attention, which occurs during the recording of eye movements, could mean that the fixations are only recorded where the eye is looking, and not what the mind is actually processing. There is therefore always the possibility that fixation and attention processes are not directly coupled, which is an inherent challenge for cognitive research (Wolff, 2015). In addition, conclusions such as that certain students are often fixated on basal bottom-up processes could be traced back. For example, dynamic stimuli (such as students in the classroom) or other factors such as certain colors automatically attract attention rather than being directly linked to top-down information (cf. Itti, 2005). The conclusion for Study 6, for example, that novices are aware of the importance of students for classroom management may therefore only be based on bottom-up processes and not on their knowledge base. It would be conceivable to conduct a study comparing eye movements in teaching

stimuli with novices from other domains and prospective teachers in order to obtain more precise information about them. Assuming that all novices would have similar fixations on the students, this would rather speak for bottom-up processes due to dynamic characteristics. If, on the other hand, the prospective teachers would fixate on the students more often than the other novices, this would suggest top-down participation already for novices. Altogether, triangulation is essential for a well-founded interpretation of eye movement data. Therefore, eye-tracking needs the use of several data sources to draw conclusions about the underlying cognitive processes or the cognitive model (Choi, Mosley, & Stark, 1995). Several of the studies dealt with this issue, as Studies 1 and 2 cross-validated their eye-tracking measures with an assessment of the lesson (with the CLASS-instrument, Pianta, La Paro, & Hamre, 2008). Studies 2 and 6 integrated additional information about the students in the classroom (in Study 2, with data of behavior or academic performance, in Study 6, with predefined roles of the simulated students). Verbal information from the participants were included in Study 9 (thinking aloud about the classroom scenes), Study 10 (to assess the awareness of the disturbing students), as well as in Studies 7 and 8 (timestamps; the participants needed to press a button when they noticed relevant scenes in the video).

Further analyses are also appropriate with regard to some parameters, such as scanpath analysis (Studies 3 and 5). These indicators offer a promising opportunity to gather additional information from the data, such as if and when teachers perceive situations and the order of fixed objects, which in turn could be used to draw conclusions about underlying cognitive processes (e.g. didactical strategies or aims).

5.5 Limitations of the systematic review

Publication bias is one of the major limitations of especially systematic literature reviews. (cf. Nelson, Simmons & Simonsohn, 2018; Zawacki-Richter et al., 2020). By using three different databases (ERIC, PsycINFO and the Web of Science), no restrictions in the search criteria (except for content) and a complementary search via Google Scholar, and additional referencing, we tried to keep the publication bias as low as possible. However, it cannot be excluded that there is further literature that was not found, because the search could not be all-encompassing.

Lastly, the present systematic review can give no information about the overall statistical effects due to the limited number of studies available and their heterogeneity concerning the sample, parameters and method of analysis. Similarly, due to the limited number of studies, it is difficult to make statements about the exact interpretation possibilities of the individual parameters. Although the few studies differed in some serious points, the results nevertheless indicate that professional vision, especially noticing, can be acquired using eye-tracking. A meta-analysis, if a majority of studies are available, would be appropriate in order to be able to really

assess the suitability of individual parameters. In addition, the other possible influencing factors such as sample or methodology would have to be controlled.

References¹

- Barth, V. L. (2017). *Professionelle Wahrnehmung von Störungen im Unterricht [Professional vision of disorders in the classroom]*. Wiesbaden: VS.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, 35(5), 463–482. doi: 10.1016/S0883-0355(02)00004-6
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift Für Psychologie*, 223(1), 3–13. doi: 10.1027/2151-2604/a000194
- Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes*. Springer Nature.
- Brophy, J. (1988). Educating teachers about managing classrooms and students. *Teaching and Teacher Education*, 4(1), 1–18. doi: 10.1016/0742-051X(88)90020-0
- Bucher, H.-J., & Schumacher, P. (2012). Aufmerksamkeit und Informationsselektion: Blickdaten als Schlüssel zur Aufmerksamkeitssteuerung. [Attention and information selection: gaze data as key to attention control]. In H.-J. Bucher & P. Schumacher (Eds.), *Interaktionale Rezeptionsforschung: Theorie und Methode der Blickaufzeichnung in der Medienforschung* (pp. 83–107). Wiesbaden: Springer VS. doi: 10.1007/978-3-531-93166-1_3.
- Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal of Teacher Education*, 39(3), 25–31. doi: 10.1177/002248718803900306
- Casale, G., Strauß, S., Hennemann, T., & König, J. (2016). Wie lässt sich Klassenführungsexpertise messen? Überprüfung eines videobasierten Erhebungsinstruments für Lehrkräfte unter Anwendung der Generalisierbarkeitstheorie [How can classroom management expertise be measured? Review of a video-based survey tool for teachers using generalizability theory]. *Empirische Sonderpädagogik*, 2, 119–139. Retrieved from https://www.psychologie-aktuell.com/fileadmin/download/esp/2-2016_20160609/ESP-2-2016_119-139.pdf
- Chi, M. T. H., & Glaser, R. (1988). Overview. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise*. Hillsdale, N.J: Lawrence Erlbaum Associates.
- Choi, Y. S., Mosley, A. D., & Stark, L. W. (1995). “Starkfest” vision and clinic science special issue: String editing analysis of human visual search. 72(7), 439–451. Retrieved from <https://insights.ovid.com/crossref?an=00006324-199507000-00003>.
- *Cortina, K. S., Miller, K. F., McKenzie, R., & Epstein, A. (2015). Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye-tracking with expert and novice teachers. *International Journal of Science and Mathematics Education*, 13(2), 389–403. doi: 10.1007/s10763-014-9610-5
- *Dessus, P., Cosnefroy, O., & Luengo, V. (2016). ‘Keep your eyes on ‘em all!’: A mobile eye-tracking analysis of teachers’ sensitivity to students. In K. Verbert, M.

1 References marked with an asterisk (*) indicate studies included in the systematic review.

- Sharples, & T. Klobučar (Eds.), *Adaptive and adaptable learning* (pp. 72–84). Switzerland: Springer International Publishing.
- Doyle, W. (1980). *Classroom management*. West Lafayette, IN: Kappa Delta Pi.
- Doyle, W. (1985). Recent research on classroom management: Implications for teacher education. *Journal of Teacher Education*, 36(3), 31–35. doi: 10.1177/002248718503600307
- Dreyfus, S. E. (2004). The five-stage model of adult skill acquisition. *Bulletin of Science, Technology & Society*, 24(3), 177–181. doi: 10.1177/0270467604264992
- Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Memory and Language*, 22(1), 75–87. doi: 10.1016/S0022-5371(83)80007-3
- Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, 36(2), 103–112. doi: 10.1207/S15326985EP3602_5
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64. doi: 10.1518/001872095779049543
- Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, 51(17), 1920–1931. doi: 10.1016/j.visres.2011.07.002
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23(4), 523–552. doi: 10.1007/s10648-011-9174-7
- Gegenfurtner, A., & Seppänen, M. (2013). Transfer of expertise: An eye tracking and think aloud study using dynamic medical visualizations. *Computers & Education*, 63, 393–403. doi: 10.1016/j.compedu.2012.12.021
- Gold, B., Hellermann, C., & Holodynski, M. (2017). Effects of video-based trainings for promoting self-efficacy in elementary classroom management. *Zeitschrift für Erziehungswissenschaft*, 20(1), 115–136. doi: 10.1007/s11618-017-0727-5
- Gold, B., & Holodynski, M. (2017). Using digital video to measure the professional vision of elementary classroom management: Test validation and methodological challenges. *Computers & Education*, 107, 13–30. doi: 10.1016/j.compedu.2016.12.012
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606–633. doi: 10.1525/aa.1994.96.3.02a00100
- Hershler, O., & Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with photographs. *Attention, Perception & Psychophysics*, 71(7), 1478–1486. doi: 10.3758/APP.71.7.1478
- Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of Intercultural Relations*, 10(3), 301–320. doi: 10.1016/0147-1767(86)90015-5
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye-tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford University Press.
- Housner, L. D., & Griffey, D. C. (1985). Teacher cognition: Differences in planning and interactive decision making between experienced and inexperienced teachers. *Research Quarterly for Exercise and Sport*, 56(1), 45–53. doi: 10.1080/02701367.1985.10608430
- Huang, R., & Li, Y. (2012). What matters most: A comparison of expert and novice teachers' noticing of mathematics classroom events. *School Science and Mathematics*, 112(7), 420–432. doi: 10.1111/j.1949-8594.2012.00161.x
- *Huang, Y. (2018). *Learning from teacher's eye movement: Expertise, subject matter and video modeling* (Doctoral dissertation, University of Michigan). Retrieved

- from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/145853/yizhen_h_1.pdf?sequence=1&isAllowed=y
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203. doi: 10.1038/35058500
- Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, 12(6), 1093–1123. doi: 10.1080/13506280444000661
- Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in educational science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, 10(3), 1–18. doi: 10.16910/jemr.10.1.3
- Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, 20(2), 146–154. doi: 10.1016/j.learninstruc.2009.02.019
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8(4), 441–480. doi: 10.1016/0010-0285(76)90015-3
- Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.
- Kasarskis, P., Stehwiën, J., Hickox, J., Aretz, A., & Wickens, C. (2001). *Comparison of expert and novice scan behaviors during VFR flight*. 6. Retrieved from <http://cite-seerx.ist.psu.edu/viewdoc/download;jsessionid=1CDB91EB26FA4E25D7351275E8353C3D?doi=10.1.1.453.1695&rep=rep1&type=pdf>
- König, J., Blömeke, S., Paine, S., Schmidt, W. H., & Hsieh, F.-J. (2011). *General pedagogical knowledge of future middle school teachers: On the complex ecology of teacher education in the United States, Germany, and Taiwan*, 62(2), 188–201.
- Kounin, J. S. (2006). *Techniken der Klassenführung [Techniques of classroom management]*. Münster: Waxmann.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820. doi: 10.1037/a0032583
- Kuperman, V., Bertram, R., & Baayen, R. H. (2008). *Morphological dynamics in compound processing*, 23(7), 1089–1132. doi: 10.1080/01690960802193688
- Lachner, A., Jarodzka, H., & Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instructional Science*, 44(3), 197–203. doi: 10.1007/s11251-016-9376-y
- Livingston, C., & Borko, H. (1989). Expert-novice differences in teaching: A cognitive analysis and implications for teacher education. *Journal of Teacher Education*, 40(4), 36–42. doi: 10.1177/002248718904000407
- *McIntyre, N. A., & Foulsham, T. (2018). Scanpath analysis of expertise and culture in teacher gaze in real-world classrooms. *Instructional Science*, 46(3), 435–455. doi: 10.1007/s11251-017-9445-x
- *McIntyre, N. A., Jarodzka, H., & Klassen, R. M. (2017). Capturing teacher priorities: Using real-world eye-tracking to investigate expert teacher priorities across two cultures. *Learning and Instruction*, 60, 215–224. doi: 10.1016/j.learninstruc.2017.12.003
- *McIntyre, N. A., Mainhard, M. T., & Klassen, R. M. (2017). Are you looking to teach? Cultural, temporal and dynamic insights into expert teacher gaze. *Learning and Instruction*, 49, 41–53. doi: 10.1016/j.learninstruc.2016.12.005
- Murray, W. S., Fischer, M. H., & Tatler, B. W. (2013). Serial and parallel processes in eye movement control: Current controversies and future directions. *The Quarterly Journal of Experimental Psychology*, 66(3), 417–428. doi: 10.1080/17470218.2012.759979
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. *Annual Review of Psychology*, 69. doi: doi.org/10.1146/annurev-psych-122216-011836

- Orquin, J. L., & Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behavior Research Methods*, *50*(4), 1645–1656. doi: 10.3758/s13428-017-0998-z
- Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206. doi: 10.1016/j.actpsy.2013.06.003
- Palmer, D. J., Stough, L. M., Burdinski, T. K., & Gonzales, M. (2005). Identifying teacher expertise: An examination of researchers' decision making. *Educational Psychologist*, *40*(1), 13–25. doi: 10.1207/s15326985ep4001_2
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system: Manual K-3*. Baltimore, MD: Paul H Brookes Publishing.
- Piwovar, V., Thiel, F., & Ophardt, D. (2013). Training in-service teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, *30*, 1–12. doi: 10.1016/j.tate.2012.09.007
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi: 10.1080/00335558008248231
- Prieto, L. P., Wen, Y., Caballero, D., Sharma, K., & Dillenbourg, P. (2014). Studying teacher cognitive load in multi-tabletop classrooms using mobile eye-tracking. *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces – ITS '14*, 339–344. doi: 10.1145/2669485.2669543
- Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*(1), 48–55. doi: 10.1111/1467-9280.00309
- Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In *Oxford Handbook on Eye Movements* (pp. 528–550). Oxford: Oxford University.
- Riedl, R., Brandstätter, E., & Roithmayr, F. (2008). Identifying decision strategies: A process- and outcome-based classification method. *Behavior Research Methods*, *40*(3), 795–807. doi: 10.3758/BRM.40.3.795
- Rötting, M. (2001). *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen. [Parameter systematics of eye and gaze movements for ergonomic investigations.]* Aachen: Shaker.
- Seidel, T., Blomberg, G., & Stürmer, K. (2010). 'OBSERVER'—Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht ["OBSERVER"—Validation of a video-based instrument for recording the professional vision of lessons]. In E. Klieme, D. Leutner, & M. Kenk (Eds.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (pp. 296–306). Retrieved from https://www.researchgate.net/publication/48524837_Observer_-_Validierung_eines_videobasierten_Instruments_zur_Erfassung_der_professionellen_Wahrnehmung_von_Unterricht_Projekt_OBSERVE
- Seidel, T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*, *51*(4), 739–771. doi: 10.3102/0002831214531321
- Sherin, M. G. (2001). Developing a professional vision of classroom events. In T. L. Wood, B. S. Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 75–93). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 383–395). New York, NY: Routledge, Taylor & Francis Group.

- Sneddon, A., Mearns, K., & Flin, R. (2013). Stress, fatigue, situation awareness and safety in offshore drilling crews. *Safety Science*, 56, 80–88. doi: 10.1016/j.ssci.2012.05.027
- Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM*, 48(1–2), 1–27. doi: 10.1007/s11858-016-0775-y
- Steffensky, M., Gold, B., Holdynski, M., & Möller, K. (2015). Professional vision of classroom management and learning support in science classrooms – Does professional vision differ across general and content-specific classroom interactions? *International Journal of Science and Mathematics Education*, 13(2), 351–368. doi: 10.1007/s10763-014-9607-0
- *Stürmer, K., Seidel, T., Müller, K., Häusler, J., & Cortina, K. S. (2017). What is in the eye of preservice teachers while instructing? An eye-tracking study about attention processes in different teaching situations. *Zeitschrift für Erziehungswissenschaft*, 20(S1), 75–92. doi: 10.1007/s11618-017-0731-9
- *van den Bogert, N. J. (2016). *On teachers' visual perception and interpretation of classroom events using eye tracking and collaborative tagging methodologies*. Technische Universiteit Eindhoven.
- *van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education*, 37, 208–216. doi: 10.1016/j.tate.2013.09.001
- van der Gijp, A., Ravesloot, C. J., Jarodzka, H., van der Schaaf, M. F., van der Schaaf, I. C., van Schaik, J. P. J., & ten Cate, Th. J. (2017). How visual search relates to visual diagnostic performance: A narrative systematic review of eye-tracking research in radiology. *Advances in Health Sciences Education*, 22(3), 765–787. doi: 10.1007/s10459-016-9698-1
- van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Guiding students' attention during example study by showing the model's eye movements. In J. Zumbach, N. Schwartz, T. Seufert, & L. Kester (Eds.), *Beyond knowledge: The legacy of competence* (pp. 189–196). doi: 10.1007/978-1-4020-8827-8_26
- van Wee, B., & Banister, D. (2016). How to write a literature review paper? *Transport Reviews*, 36(2), 278–288. doi: 10.1080/01441647.2015.1065456
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249–294. doi: 10.3102/00346543063003249
- Wolff, C. E. (2015). *Revisiting 'whititness': Differences in teachers' representations, perceptions, and interpretations of classroom management* (Doctoral dissertation). Heerlen, the Netherlands.
- *Wolff, C. E., Jarodzka, H., van den Bogert, N., & Boshuizen, H. P. A. (2016). Teacher vision: Expert and novice teachers' perception of problematic classroom management scenes. *Instructional Science*, 44(3), 243–265. doi: 10.1007/s11251-016-9367-z
- Wolff, C. E., van den Bogert, N., Jarodzka, H., & Boshuizen, H. P. A. (2015). Keeping an eye on learning: Differences between expert and novice teachers' representations of classroom management events. *Journal of Teacher Education*, 66(1), 68–85. doi: 10.1177/0022487114549810
- *Yamamoto, T., & Imai-Matsumura, K. (2013). Teachers' gaze and awareness of students' behavior: Using an eye tracker. *Comprehensive Psychology*, 2, 01.IT.2.6. doi: 10.2466/01.IT.2.6
- Yarbus, A. (1967). *Eye movements and vision*. New York, NY: Plenum.
- Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (2020). *Systematic Reviews in Educational Research*. Wiesbaden: VS.

Appendix

Figure A: Flow diagram of the study selection process

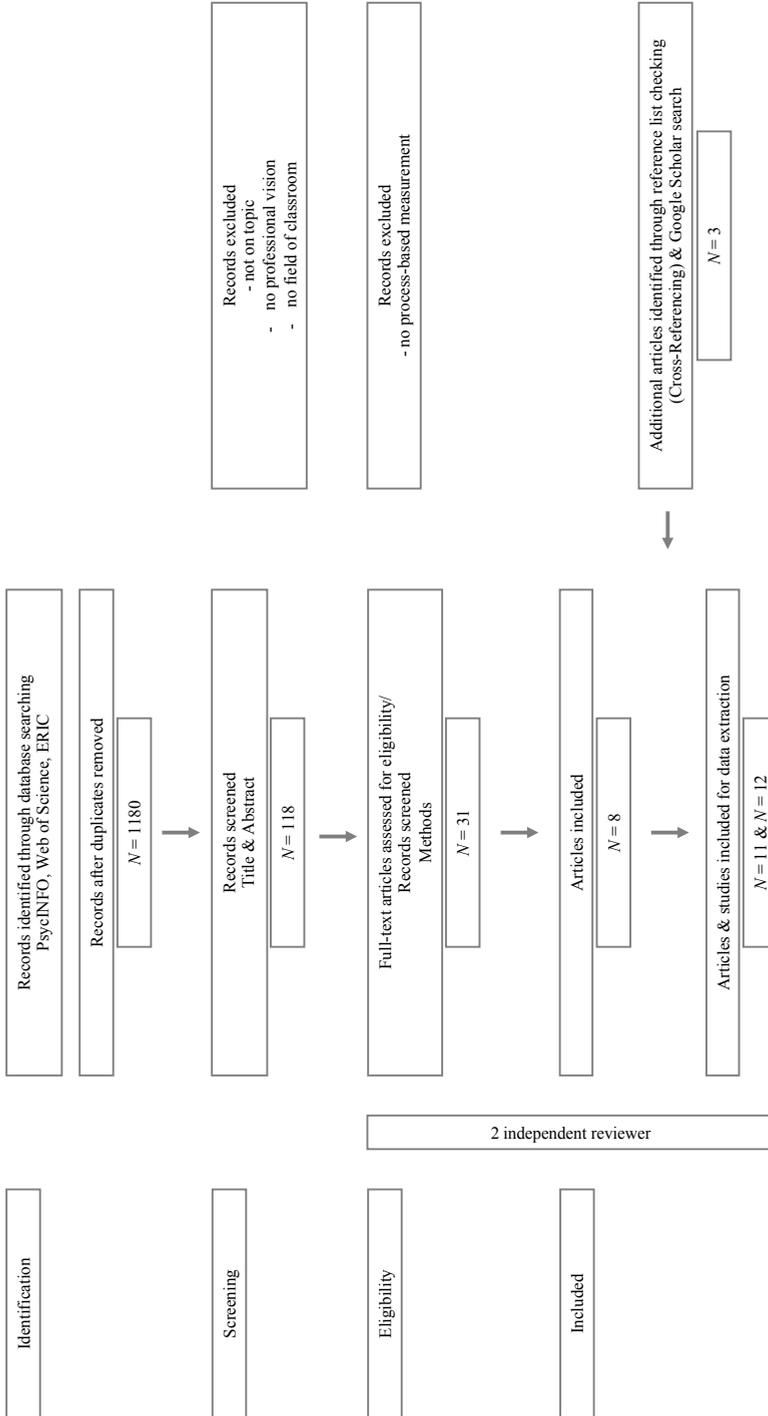


Table A: Search strings

| Topic and cluster | Search terms |
|-------------------------------|--|
| Process-based measurement | eyetrack* OR eye-track* OR eye track* OR tag* OR annotation* OR gaze* OR "eye tracking" OR tag* OR "video annotations" OR "mobile eye-tracking" OR "eye-tracking" or gaze* |
| Professional vision | "professional vision" OR "professional development" OR awareness |
| (Prospective) teachers | teach* OR expert* "teacher expertise" OR "teacher noticing" OR "teacher education" OR "teacher knowledge" OR "teacher attention" OR "expert-novice paradigm" |
| Field of classroom management | classroom* OR "classroom management" OR "classroom perception" OR "classroom simulation" OR "classroom observational assessment" OR "classroom techniques" |

Lars Burghardt, Anja Linberg, Simone Lehl, Kira Konrad-Ristau

The relevance of the early years home and institutional learning environments for early mathematical competencies

Abstract

Mathematical competencies are a core prerequisite for educational success. In the present study, we therefore examine the relevance of the early years home and institutional learning environment at the age of two years on mathematical competencies at the age of four, while controlling for the later years home learning environment using the data of the starting cohort one of the NEPS study (N = 1170). Results show positive effects of global processual aspects of the home learning environment, conceptualized as the frequency of joint activities at home and domain-specific aspects, which are mathematical and language stimulation in parent-child interactions. The effects of parent-child interactions in the early years remains when control variables and joint activities at later ages are added. Processual aspects of institutional child care were analyzed on a subsample (N = 230) and showed only small associations with later mathematical competencies.

Keywords

Home learning environment; Institutional learning environment; Mathematical competencies; Parent-child interaction; Joint activities; Early childhood education and care

Dr. Lars Burghardt (corresponding author), Chair of Early Childhood Education and Care, University of Bamberg, Markusstraße 8a, 96047 Bamberg, Germany
e-mail: lars.burghardt@uni-bamberg.de

Dr. Anja Linberg, German Youth Institute, Nockherstraße 2, 81541 Munich, Germany
e-mail: linberg@dji.de

Dr. Simone Lehl · Kira Konrad-Ristau, M.Sc., Department Psychology I – Developmental Psychology, University of Bamberg, Markusplatz 3, 96047 Bamberg, Germany
e-mail: simone.lehl@uni-bamberg.de
kira.konrad-ristau@uni-bamberg.de

Die Bedeutung frühkindlicher häuslicher und institutioneller Lernumwelten für frühe mathematische Kompetenzen

Zusammenfassung

Mathematische Kompetenzen sind eine Grundvoraussetzung für den Bildungserfolg. In der vorliegenden Studie untersuchen wir daher die Relevanz der häuslichen und institutionellen Lernumgebung von Kindern im Alter von zwei Jahren für mathematische Kompetenzen im Alter von vier Jahren anhand Daten der Startkohorte 1 des NEPS (N = 1170 Kinder) und kontrollieren die häusliche Lernumgebung zu späteren Zeitpunkten. Die Ergebnisse zeigen positive Effekte globaler prozessualer Merkmale der häuslichen Lernumgebung, wie der Häufigkeit gemeinsamer Aktivitäten zu Hause, sowie domänenspezifische Effekte der mathematischen und sprachlichen Anregung in Eltern-Kind-Interaktionen. Die Auswirkungen der frühen Eltern-Kind-Interaktion bleiben bestehen, wenn Kontrollvariablen und gemeinsame Aktivitäten aus späteren Wellen ergänzt werden. In einer Teilstichprobe (N = 230) wurden Aspekte der institutionellen Betreuung analysiert, hier zeigen sich nur geringe Zusammenhänge bezogen auf die späteren mathematischen Kompetenzen.

Schlagworte

Häusliche Lernumwelt; Institutionelle Lernumwelt; Mathematische Kompetenzen; Eltern-Kind-Interaktion; Gemeinsame Aktivitäten; Frühkindliche Erziehung Bildung und Betreuung

1. Introduction

The stimulation children experience in different learning environments are jointly responsible for the formation of inter-individual differences in various areas of competencies. In the early childhood phase, the family and institutional child care can be regarded as the central learning environments. Even though research shows that both environments play an important role in fostering competencies and for the general upbringing of children (for an overview: Lehl, 2018; Melhuish et al., 2015), comparatively little is known about the effect of the home learning environment (HLE) and the institutional learning environment of early childhood education and care settings (ECEC) for children under the age of three. The focus on the very early years of a child's life seems essential not only for general child development but also for the development of mathematical competencies. First learning experiences, such as early numerical or mathematical activities like counting, comparing, or sorting, take place long before school entry. Accordingly, difficulties in dealing with mathematical phenomena are considered a cumulative process that can begin in the first years of life (Geary, Hamson, & Hoard, 2000). Mathematical

competencies in particular are essential both for later education and for participation in social life (Duncan et al., 2007). Studies demonstrate that these early mathematical competencies, in the age range of three to six years, predict quite well the development of later mathematical competencies and also general school success, even when the intelligence of the child is taken into account (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Dornheim, 2008; Jordan, Kaplan, Locuniak, & Ramineni, 2007). The foundations for these competencies are already laid during the first four years of a child's life (Feigenson, Dehaene, & Spelke, 2004; Carey, 2009; Spelke, 2017). Based on inherent systems, such as the nonverbal systems of quantity representation (Feigenson et al., 2004), different mathematical skills develop further or emerge, like the perception of quantities, counting, or the idea of space and time (Krajewski, Grüßing, & Peter-Koop, 2009). The research landscape for the German early childhood educational system regarding very young children is characterized mostly by gaps. Additionally, conclusions from existing international studies have to be drawn from with caution because there are huge differences in maternal leave (International Labour Office, 2013), day care regulations (Waldfogel, 2001), and the early child care system itself (OECD, 2013). From a psychological-educational point of view, the emergence of the precursors of mathematical competencies and how they can be supported from early on are major research interests. Especially numerical or mathematical activities like counting, comparing, or sorting are supposed to foster the development of these competencies.

The present paper uses data from the starting cohort one of the German National Education Panel Study (NEPS; Blossfeld, Roßbach, & von Maurice, 2011) to analyze the relation between the experiences children have in their families as well as in institutional childcare under the age of three, and mathematical competencies at the age of four. The starting cohort one began as a nationally representative sample of 3481 children and their families in the year 2012. This period is of particular interest since daycare regulations, especially those for children under the age of three, have changed fundamentally. For example, since 2013, each child older than one year is entitled to a place in a childcare center or in family-based daycare. While only 15.7 % of German children attended institutional childcare in 2007, the attendance rate increased to 34.3 % in 2019 (German Federal Statistical Office, 2019). Thus, the NEPS includes one of the first cohorts of children and their parents exposed to the changed daycare situation in Germany.

For a better understanding, a short overview of the German ECEC-system will be provided. Usually, the German ECEC system is divided into two institutional settings: one for children aged 0 to 3 years and one for those aged 3 to 6. Besides this division, there is a mixed form, where all children age 0 to 6 are cared for. Note that these are age-mixed groups within the particular age range. Since 1996, each child aged 3 years or older has been entitled to a place in a German child day care centre. In 2013, this legal right was expanded to entitle each child older than 1 year to a place in a child care centre or in family-based day care. With an attendance rate for 3- to 6-year-olds of 93.0 %, almost every child in this age group has some ECEC experience, and the use of such services is highly accepted. The rate for

the under-3s is 34.3 % (1.9 % for 0- to 1-year-olds, 37.1 % for 1- to 2-year-olds and 63.2 % for the 2- to 3-year-olds; German Federal Statistical Office, 2019). The lower attendance rates for the under-3s are partially due to the laws on parental benefits and parental leave. If parents decide to stay at home and care for their child after birth, the government will replace their income loss by 65 to 100 % for a maximum of 14 months. Since July 2015, an additional law called ‘ElterngeldPlus’ even allows a further 4 months of combining parental allowance and part-time work (Federal Ministry for Family Affairs, Senior Citizens, Women and Youth, 2015).

2. Theoretical framework and state of research

2.1 Conceptualizing early years learning environments

According to bio-ecological development theories (Bronfenbrenner & Morris, 2006), proximal processes, i.e. activities and interactions of the individual with his/her environment, are central to the acquisition of mathematical competencies (as well as other competencies). The research further differentiates these processes into domain-specific and global processes (Kluczniok, Lehl, Kuger, & Roßbach, 2013; Lehl, 2018). Domain-specific processes, like reading to the child or playing number board games, are supposed to stimulate specific developmental areas, such as language or mathematics (e.g. LeFevre, Polyzoi, Skwarchuk, Fast, & Sowinski, 2010; Lehl, Ebert, Blaurock, Roßbach, & Weinert, 2020). Global processes, on the other hand, are defined as activities or interactions that cannot be assigned to specific developmental areas or in which individual content areas are classified in a global picture of the respective processes taking place (e.g., painting or pretend play; Lehl, 2018). Another theoretical strand is represented by the interactionist assumptions of Vygotski, which stress the importance of interactions of the child with its immediate environment. To a certain extent, children acquire competences by dealing with the environment in which they grow up. Especially with regard to specific, culturally charged knowledge (like mathematics), the acquisition of knowledge requires mediation in co-constructive social interactions. Referring to the “zone of proximal development” which can be understood as the distance between the current child’s stage of development and the potential stage of development, the interaction partners (here the parents) are of particular importance as they represent the knowledge before it can be internalized by the children (Leseman & de Jong, 1998). According to this, higher mental functions are acquired through the support of a more competent partner, by connecting to the current level of development and helping the learner to grow beyond his own abilities, which e.g., occurs in joint activities. Similar to this concept, Wood, Wood, and Middleton (1978) have already proposed the term “scaffolding” with reference to Vygotski’s theory. Scaffolding refers to those strategies of the more competent

partner that ensure a sensitive provision and sensitive withdrawal of support during a joint problem-solving process (Linberg, 2018).

In the period of early childhood, the family is certainly the first and central learning environment (Bronfenbrenner & Morris, 2006). With increasing age, other learning environments also become significant. In early childhood, institutional child care settings often represent another learning environment. Since the introduction of a legal entitlement to a place in child care from the age of one and the associated expansion, the care rate for children under three years has more than doubled in the last ten years in Germany. In this learning environment, mathematical support has also been established as a clear area of responsibility for educational work in the early years (Youth Ministers' Conference & Culture Ministers' Conference, 2004) and is anchored in most educational plans of the federal states.

2.2 State of research

2.2.1 Relationships between home learning environment and mathematical competencies

The term “home learning environment” comprises several aspects of the frequency and quality of stimulating activities between children and their primary caregivers as well as the availability of learning materials at home (Linberg, Lehl, & Weinert, 2020; see Bradley & Corwyn, 2002 for review). Such proximal home learning experiences have been shown to be associated with mathematical competencies. For example, the global processes (e.g., parental responsiveness) that the child experiences at the age of three years or older are predictive of later mathematical abilities (Burchinal, Peisner-Feinberg, Pianta, & Howes, 2002; Bradley, Corwyn, McAdoo, & Coll, 2001). Results of the Effective Provision of Preschool Education (EPPE) study show long-term effects of global processes (measured as educationally relevant joint activities with the child) at the age of three years for mathematical competencies at the age of seven, and even showing effects up to 16 years (Melhuish et al., 2008; Sylva, Melhuish, Sammons, Siraj, & Taggart, 2014).

However, it is often pointed out that effects must be examined domain-specifically (Lehl, et al., 2020; Sénéchal & LeFevre, 2002). Math-specific processes, such as engaging children in activities like counting, learning numbers and comparing sizes have proven to be important for the development of math competencies (e.g., Elliott & Bachman, 2018; Niklas & Schneider, 2012; Susperreguy, Di Lonardo Burr, Xu, Douglas, & LeFevre, 2020; Zippert & Rittle-Johnson, 2020). Studies also point to cross-domain effects, meaning that not only math-specific, but also language-specific processes are connected to math competencies (sometimes they are even more pronounced: Lehl et al., 2020; LeFevre, Polyzoi et al., 2010; Napoli & Purpura, 2018). This might be the case because of the close relationship of domain-specific processes in the home learning environment, or it might be a result of the use of mathematical language (“Math talk”) taking place during literacy ac-

tivities (Lehl, 2018; Lehl et al., 2020) which is also connected to the development of mathematic competencies (Ramani, Rowe, Eason, & Leech, 2015; Eason & Ramani, 2018). However, a large body of research on the importance of (global and domain-specific) processes in the home learning environment focuses on children from three years onwards. Studies focusing on the importance of the home learning environment for children under the age of three are sparse and mainly focused on social and/or language outcomes (e.g., Lehl, 2018; Mistry, Benner, Biesanz, Clark, & Howes, 2010; Rodriguez et al., 2009). One assumption is that the effects of the HLE during the preschool years might be an effect of former HLE-experiences in the toddler phase. Tamis-LeMonda, Luo, McFadden, Bandek, and Vallatton (2019) investigated the unique relation between quality of interactions and stimulating activities during the ages of one to three years on academic achievement in class five. They showed that early HLE effects were mainly mediated through its association with pre-academic skills and to a small degree (9 % of the variance) through its association with grade 5 HLE (Tamis-LeMonda et al., 2019). However, no mediating effects of consecutive HLE were found by Rodriguez et al. (2009) who investigated the effects of the home literacy environment at age 14, 24 and 36 months on children's language and literacy outcomes. They reported unique contributions of each HLE measure at each age for vocabulary development. Similarly, although focusing on preschool aged children, Lehl et al. (2020) found that preschool HLE effects on reading and mathematical competencies at age 12 are mediated mainly through children's competencies at pre- and primary school, and only partly through primary school HLE. Likewise, Sammons et al. (2015) reported that primary school HLE adds to the effects of preschool HLE measures on child outcomes at age 16. Thus, regarding the addition or mediation of effects of consecutive HLE measures the findings are mixed.

2.2.2 Relationships between ECEC and mathematical competencies

A dramatic increase in childcare attendance in Germany has motivated our study to investigate whether the variations in attendance duration and the variation in indicators for process quality in the daycare-centers contribute to children's early differences in mathematical competencies. Referring to the structure-process model of quality (Bryant, Burchinal, & Zaslow, 2011; NICHD ECCRN, 2002a) we understand process quality as an umbrella term for all interaction processes between children and their educators (e.g., games, activities like shared picture book reading or dealing with numbers), their peers as well as the spatial-material environment.

Studies reveal positive effects of overall preschool attendance and the quality of the preschool for children's social and cognitive development, including mathematic knowledge (for an overview: Ulferts & Anders, 2016). Research from the German Studies Educational Processes, Competence Development and Selection

Decisions at Pre- and Elementary School Age (BiKS) and Preschool of the Future in Bavaria (KidZ) reveal that global but especially domain-specific processes children experienced in child care are important for explaining differences in early and later math development (Anders et al., 2012; Lehl, Kluczniok, & Roßbach, 2016; Lehl, Kluczniok, Roßbach, & Anders, 2017). Furthermore, early global and specific math intervention programs with preschoolers proved to be highly effective in terms of children's mathematical development (Clements & Sarama, 2007; Skillen, Berner, & Seitz-Stein, 2018; Hauser, Vogt, Stebler, & Rechsteiner, 2014). Again, results for children under the age of three are sparse: Loeb, Fuller, Kagan, and Carrol (2004) find that the attendance of child care at the age of two to three years is positively associated with the development of mathematical competencies at the age of five. Similarly, results from the Effective Provision of Preschool Education (EPPE) study also showed positive effects of attending child care before the age of 3 (e.g., Sylva, Stein, Leach, Barnes, & Malmberg, 2011). Wylie, Ferral, Hodgen, Thompson, and New Zealand Council for Educational Research (2006) point out that the overall duration of early child care is still connected to math skills at the age of 14 years. In a small Austrian sample Baumeister, Rindermann, and Barnett (2014) also found positive effects of crèche attendance on children's cognitive outcomes. However, Tietze et al. (2013), referring to a German sample, detected no attendance effects but did detect quality effects: Children attending high quality child care showed higher cognitive skills. Results from the National Institute of Child Health and Human Development Study of Early Child Care and Youth Development (NICHD SECCYD) also point in this direction: They measured quality in child care when children were six, 15, 24, 36, and 54 months old (NICHD ECCRN, 2002b). The results reveal that children receiving higher quality child care even under the age of three show better cognitive and social outcomes, including math outcomes at the age of 54 months (NICHD ECCRN, 2006; NICHD ECCRN & Duncan, 2003). Thus, whether attending a child care center at toddler age per se benefits children's development is not clearly documented (Jaffee, Van Hulle, & Rodgers, 2011), and might be moderated by child background variables (e.g., migrations background; Klein & Sonntag, 2017). However, the quality children experience in child care has consistently been shown to boost their development (Melhuish et al., 2015).

3. The present study

Only a few large-scale longitudinal studies directly assess both child development and HLE in the very early periods of a child's life. Developmental progress and child education do not result from experiences in one single learning environment. Applying this model, depicted by Bronfenbrenner and Morris (2006) to the very first years of life, a question arises regarding how HLE and ECEC are connected to children's competencies.

In particular, we address the following research questions:

- 1) Are there connections between global and domain-specific home learning environment at the age of two years and the time spent in ECEC under the age of three with mathematical competencies at the age of four years?
- 2) Does the effect of early HLE persist, when later HLE experiences at age 36 and 48 months are controlled for?
- 3) Within the group of children attending childcare under the age of three: Does the frequency of educational activities in ECEC add to the effects of HLE and ECEC duration on mathematical competencies at the age of four years?

The article is based on the following hypothesis: the early home learning environment children grow up in, plays an important role in fostering their competencies. For mathematical competencies, we suspect that especially domain-specific processes in the field of numeracy as well as the general HLE are connected to the child's mathematical competence. Regarding the addition or mediation of early and later HLE effects, the mixed research findings induce us to assume that effects of the early years home learning environment are at least partly mediated by later home learning environment.

With regard to the effects of institutional child care, we expect both positive effects for time spent in ECEC as well as extent of activities, as both can be seen as an enrichment in addition to the impact of the family. As described in the theoretical framework and the state of research, we suggest that the effect of institutional ECEC is smaller than the effect of the family a child grows up in, as it is the first and central learning environment.

4. Method

4.1 Sample

As already mentioned, the study draws on data of the Starting Cohort One of the German National Education Panel Study (NEPS¹; Blossfeld & Roßbach, 2019), which started as a nationally representative sample of 3481 children born in 2012 and their families. From 2013 onwards, each child older than one year was entitled to a place in a child care center or in family-based day care, making this NEPS cohort one of the first to include children and their parents who have been exposed to the changed day care situation in Germany. We use the first five waves of this

1 This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Newborns, doi:10.5157/NEPS:SC1:6.0.0. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg in cooperation with a nationwide network.

cohort, in which the children were 7 months of age (wave 1), 13 months (wave 2), 27 months (wave 3), 38 months (wave 4), and 48 months (wave 5).

Information was gathered via computer-assisted interviews, videotaped and coded interactions of parents and their children, as well as competence tests of the children. By including data of the home learning environment, parent-child interactions and variables of institutional child care in wave three, it is possible to have information on different learning experiences in both central learning environments for children under the age of three. The panel design allows adding data of the HLE from later waves to prove whether the effects of early HLE remain. The mathematical competencies measured in the fifth wave of the NEPS were used in the study (children's age: 48 months).

Due to missing data in the central variables and drop-out over the five waves, the sample for the first research question is $N = 1170$. For the second research question, which examines activities in ECEC our sample size is reduced to $N = 230$. Note that the comparatively small sample size only includes children who were being cared for at wave three and the initial sample of the educators, which gives information on institutional child care from 560 educators, as information on ECEC activities were gained by a drop-off questionnaire for educators, which had a return rate of 31.6 %.

4.2 Measures

4.2.1 Dependent variable

Mathematical competencies of the children were assessed by using the standardized test “KiKi – Kieler Kindertagertest” (Grüßing et al., 2013). This test contains tasks referring to subareas like ‘sets, numbers and operations’, ‘units and measuring’ or ‘space and shapes’ (for more information on the theoretical background or validity-analyses: Neumann et al., 2013; Knopp et al., 2014). We used the weighted maximum likelihood estimates in our analysis ($M = .08$, $SD = 1.00$, Table 1).

4.2.2 Independent variables

Home learning environment (HLE). For information on the HLE we used two sets of different indicators (Table 1). One scale was formed to give information about *frequency of joint activities at home*. These indicators include activities like ‘shared picture book reading’ or ‘dealing with numbers’ (0 = never; 7 = several times a day). The indicator was computed for wave three (Cronbach's $\alpha = .53$, seven items, $M = 4.54$, $SD = .83$), four (Cronbach's $\alpha = .55$, six items, $M = 4.36$, $SD = .89$) and five (Cronbach's $\alpha = .59$, six items, $M = 4.21$, $SD = .91$).

The second set of indicators gives information about parental stimulation in a semi-standardized play situation, which was videotaped and coded. The play sit-

uation (with a standardized toy set) lasted ten minutes and was conducted in the third wave at the home of the family. Trained coders using a system adapted from the NICHD study (NICHD Early Child Care Research Network, 1999) rated the parent-child interaction on a five point scale, indicating whether the described behavior was (1) not at all to (5) highly characteristic for the parent (see Linberg, Mann, Attig, Vogel, Weinert, & Roßbach, 2019 for a description of the procedure and instrument). As we focus on domain-specific stimulation, we used two variables in our analysis: *parent-child interaction: numeracy stimulation* ($M = 1.43$, $SD = .60$) and *parent-child interaction: language stimulation* ($M = 3.41$, $SD = .75$). Language stimulation “captures the amount and quality of verbal enrichment of the play situation, including prompting and expanding child’s verbalizations, asking open ended questions, correcting mistakes, decontextualization / verbal distancing of subject matters. Numeracy stimulation covers inclusion of mathematical concepts in the play situation at a very basic level, e.g. counting, comparing, sorting, distinguishing patterns, in amount and variation” (Linberg et al., 2019, p. 7). Inter-rater reliability was checked by double-coding 20 % of randomly selected videos (inter-rater agreement for both scales > 90 %).

ECEC. For institutional child care we used the time a child spent in ECEC in his/her first three years (*duration* in months, $M = 12.34$, $SD = 9.23$, Table 1). Additionally, for the subsample we computed two variables, which give information about the frequency of activities as indicators for process quality. Here we followed a study of Linberg, Kluczniok, Burghardt, and Freund (2017) where observed quality (measured with the German version of the Infant/Toddler Environment Rating Scale, Harms, Cryer & Clifford, 2007) was linked to the NEPS-questionnaire. The goal of the study was to identify a small set of staff reported variables, which are linked to observed quality. For activities, two scales (0 = never; 7 = several times a day) were generated from data of the third wave: *Everyday activities* which include ‘reading out loud/looking at picture books’, ‘making music/singing/dancing’, ‘playing with puppets and stuffed animals’ and ‘building blocks or stacking games’ (Cronbach’s $\alpha = .70$, $M = 6.15$, $SD = .85$) and *extensive activities* which consists of three variables demanding more time and effort: ‘moving outdoors’, ‘tinkering, painting, kneading’ and ‘playing in sand or in/with water’ (Cronbach’s $\alpha = .57$, $M = 5.54$, $SD = .87$, Table 1). Linberg et al. (2017) showed that the extensive activities are especially linked to observed quality.

4.2.3 Control variables

As depicted in the theoretical model, it is important to investigate the effects of different learning environments, while considering the individual characteristics of the child, respective of the family. This seems especially important for children under the age of three in the context of German institutional child care. With regard to the theoretical approach of the economic consumer choice model, which mainly refers to the trade-off of returning to work through using child care (Boudon,

1974), as well as the state of research which shows that access to institutional child care is not accessible for all families in the same way (Burghardt, 2019), meaning that higher educated parents are more likely to return (earlier) to work and parents with a migration background are less likely, the following control variables (see table 1) are added to reduce the effects of selection into child care: migration background of the family (based on data on the first language spoken in the home, 0 = no migration background; 1 = migration background, 13 %) and mother’s educational status (ISEI 08; Ganzeboom, 2010; $M = 60.67$, $SD = 18.27$). For the individual characteristics of the child, we used the child’s age (at wave five, $M = 50$, $SD = 1.56$) and the child’s gender (0 = girl; 1 = boy, 50 %).

Table 1: Descriptives

| | <i>N</i> | <i>M</i> | <i>SD</i> | Min | Max |
|--|----------|----------|-----------|-------|-------|
| Mathematical competencies (wave 5) | 1170 | .08 | 1.00 | -3.46 | 3.19 |
| HLE (wave 3) | | | | | |
| Joint activities at home (0=never; 7=several times a day) | 1170 | 4.54 | .83 | 1.43 | 6.57 |
| Parent-child interaction: mathematical stimulation (1=not at all characteristic; 5=very characteristic) | 1170 | 1.43 | .60 | 1 | 5 |
| Parent-child interaction: language stimulation (1=not at all characteristic; 5=very characteristic) | 1170 | 3.41 | .75 | 1 | 4 |
| ECEC | | | | | |
| Duration of time spent under the age of three (in months) | 1170 | 12.34 | 9.23 | 0 | 30 |
| Everyday activities (wave 3) (0=never; 7=several times a day) | 230 | 6.15 | .85 | 3.25 | 7 |
| Extensive activities (wave 3) (0=never; 7=several times a day) | 230 | 5.54 | .87 | 2.00 | 7 |
| Control variables: child and family | | | | | |
| Child’s age at wave 5 (in months) | 1170 | 50.00 | 1.56 | 46 | 54 |
| Child’s gender (0=girl, 1=boy) | 1170 | .50 | .50 | 0 | 1 |
| Migration background (0=no MB, 1=MB) | 1170 | .13 | .34 | 0 | 1 |
| Mother’s educational status (ISEI 08) | 1170 | 60.67 | 18.27 | 11.74 | 88.96 |
| Control variables: HLE at wave 4 & 5 | | | | | |
| Joint activities at home wave 4 (0=never; 7=several times a day) | 1170 | 4.36 | .89 | 1.67 | 6.5 |
| Joint activities at home wave 5 (0=never; 7=several times a day) | 1170 | 4.21 | .91 | 0.83 | 6.5 |

4.3 Statistical analyses

To examine the impact of the early years home and institutional learning environment on children's mathematical competencies, a stepwise procedure was used. First, linear regression analyses were run where in Model 1 mathematical competencies were regressed on frequencies of joint activities at home as an indicator for global HLE. In Model 2, mathematical competencies were in addition to Model 1 regressed on the two domain-specific HLE indicators of parent-child interaction 'mathematical stimulation' and 'language stimulation' (Model 2). Third, in order to show if there is an effect of time spent in ECEC, the variable for duration was included. In the fourth model, control variables were added. To indicate whether the effect of the early years HLE (wave 3) and time spent in institutional child care under the age of three persist, we include joint activities at home from wave four and five (Model 5). To answer the last research question, we take a detailed look at those children who attended ECEC and analyze the additional effect of activities on mathematical competencies. Again, the first model includes the indicator for global HLE, the second model includes domain-specific HLE. The third model includes the duration and the scales for activities in institutional child care, before the control variables are added in the last two models. The analysis has been carried out using IBM SPSS 25.

5. Results

5.1 Relation between the early years HLE and time spent in ECEC with later mathematical competencies

The results for the effect of the early years home learning environment on mathematical competencies are shown in Table 2. Model one shows that there is a positive relationship between frequencies of joint activities at home and the outcome variable. However, the explained variance is quite small ($\text{adj. } R^2 = .01$). The second model reveals a positive effect of domain-specific HLE – language stimulation during parent-child interaction – on mathematical competencies. The time children spent in institutional child care reveals a positive tendency, meaning more time spent in ECEC is connected to higher mathematical competencies (Model 3). Note that the mathematical stimulation in parent-child interaction also reveals a positive tendency. By adding individual characteristics of the children and their families (Model 4), the explained variance increases to 11 %. Older children as well as children whose mother has a higher educational status reveal higher mathematical competencies. There is also a gender effect, displaying higher competencies in girls than in boys. At a significance level of $p < .10$, there is also an effect demonstrating lower mathematical skills in children with a migration background. Note that the effect of ECEC duration vanishes and the effects of joint HLE activities and lan-

Table 2: Regression of early years HLE and time spent in ECEC on later mathematical competencies

| Mathematical competencies at wave 5 (N = 1170) | | | | | | | | | | |
|--|--------------|-----|---------------|-----|------------------|-----|-------------------|-----|-------------------|-----|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| | β | SE | β | SE | β | SE | β | SE | β | SE |
| Step 1: global HLE (wave 3) | | | | | | | | | | |
| Joint activities at home (0=never; 7=several times a day) | .09** | .04 | .08** | .04 | .09 ⁺ | .04 | .06 ⁺ | .03 | .01 | .04 |
| Step 2: domain-specific HLE (wave 3) | | | | | | | | | | |
| Parent-child interaction: mathematical stimulation (1=not at all characteristic; 5=very characteristic) | | | .05 | .05 | .05 ⁺ | .05 | .07* | .05 | .07* | .05 |
| Parent-child interaction: language stimulation (1=not at all characteristic; 5=very characteristic) | | | .13*** | .04 | .13*** | .04 | .08** | .04 | .08** | .04 |
| Step 2: ECEC | | | | | | | | | | |
| Duration of time spent under the age of three (in months) | | | | | .05 ⁺ | .00 | .02 | .00 | .03 | .00 |
| Step 3: Control variables: child and family | | | | | | | | | | |
| Child's age (in months) | | | | | | | .22*** | .02 | .22*** | .02 |
| Child's gender (0=girl, 1=boy) | | | | | | | -.07** | .06 | -.07* | .06 |
| Migration background (0=no MB, 1=MB) | | | | | | | -.05 ⁺ | .08 | -.05 ⁺ | .08 |
| Mother's educational status (ISEI 08) | | | | | | | .18*** | .00 | .18*** | .00 |
| Step 4: Control variables: HLE at wave 4 & 5 | | | | | | | | | | |
| Joint activities at home wave 4 (0=never; 7=several times a day) | | | | | | | | | .06 ⁺ | .04 |
| Joint activities at home wave 5 (0=never; 7=several times a day) | | | | | | | | | .03 | .04 |
| R² | .01** | | .03*** | | .03*** | | .12*** | | .12*** | |
| adj. R² | .01 | | .03 | | .03 | | .11 | | .12 | |

+p < .10. *p < .05. **p < .01. ***p < .001. Significance of the ANOVA is indicated at R².

guage stimulation at home decrease while the effect of mathematical stimulation at home increases when adding the family background characteristics. In the final model (Model 5), the effects at the individual level remain the same. Even after controlling for joint HLE activities in the fourth and fifth wave, there are still significant effects of early parent-child interactions on mathematical competencies when the child was two years old but there are no effects for joint-activities at this early age. Overall, the explained variance rises to 12 %.

5.2 Relation between early years HLE, time spent and activities in ECEC on later mathematical competencies

To analyze the second research question, we used a subsample of our initial sample, which gives information about the group of children attending ECEC and the activities carried out in institutional child care at the age of two years. Like the analyses on the full sample, the first model includes HLE indicators, the second model includes the duration and the scales for activities in institutional child care, and in the final two models, the control variables were added (Table 3). In the comparatively small subsample, we do find a small tendency for global HLE (Model 1). In the second model, we do not find effects of either global or domain-specific HLE. The same is true when we add the duration and activities in ECEC (Model 3). When control variables are included, a positive tendency for mathematical stimulation in parent-child interaction ($\beta = .12$; $p < .10$) and for extensive activities in ECEC ($\beta = .12$; $p < .10$) arises. This model explains 6 % of the variance. Adding joint activities at home from wave four and five does not change the picture. All coefficients remain unchanged with the exception of mathematical stimulation, which decreases from $\beta = .12$ to $\beta = .11$.

6. Discussion

Based on the assumptions of ecological interactionist theory (Bronfenbrenner & Morris, 2006; Vygotski & Cole, 1978), our study explored the relevance of the early years home and institutional learning environments for early mathematical competencies. To get a better understanding of the results and also their validity, it seems necessary to point out the following limitations:

We described above, that we are using longitudinal data of the NEPS. As this is true for variables like 'joint activities at home' or 'duration of time spent in ECEC,' we could only refer to the mathematical competencies of the children in the fifth wave (age 4). As this is the first time it was assessed in the starting cohort one of the NEPS study, we were not able to analyze developmental progress but developmental status. However, we could not take previous competencies into account

Table 3: Regression of early years HLE and time spent in and activities in ECEC on later mathematical competencies

| | Mathematical competencies at wave 5 (N = 230) | | | | | | | | | |
|---|---|-----|------------|-----|------------|-----|------------------------|-----|------------------------|-----|
| | Model 1 | | Model 2 | | Model 3 | | Model 3 | | Model 4 | |
| | β | SE | β | SE | β | SE | β | SE | β | SE |
| Step 1: global HLE (wave 3) | | | | | | | | | | |
| Joint activities at home (0=never; 7=several times a day) | .11 ⁺ | .09 | .10 | .09 | .11 | .04 | .10 | .09 | .06 | .10 |
| Step 2: domain-specific HLE (wave 3) | | | | | | | | | | |
| Parent-child interaction: mathematical stimulation (1=not at all characteristic; 5=very characteristic) | | | .09 | .13 | .10 | .05 | .12 ⁺ | .13 | .11 ⁺ | .13 |
| Parent-child interaction: language stimulation (1=not at all characteristic; 5=very characteristic) | | | .04 | .10 | .04 | .04 | -.01 | .10 | .00 | .10 |
| Step 3: ECEC | | | | | | | | | | |
| Duration of time spent under the age of three (in months) | | | | | -.04 | .01 | -.06 | .01 | -.07 | .01 |
| Everyday activities (wave 3) (0=never; 7=several times a day) | | | | | -.09 | .09 | -.10 | .09 | -.11 | .09 |
| Extensive activities (wave 3) (0=never; 7=several times a day) | | | | | .12 | .08 | .12 ⁺ | .08 | .12 ⁺ | .08 |
| Step 4: Control variables: child and family | | | | | | | | | | |
| Child's age (in months) | | | | | | | .19 ^{**} | .04 | .19 ^{**} | .04 |
| Child's gender (0=girl, 1=boy) | | | | | | | -.03 | .14 | -.02 | .14 |
| Migration background (0=no MB, 1=MB) | | | | | | | -.03 | .22 | -.03 | .22 |
| Mother's educational status (ISEI 08) | | | | | | | .15 [*] | .00 | -.15 [*] | .00 |
| Step 5: Control variables: HLE at wave 4 & 5 | | | | | | | | | | |
| Joint activities at home wave 4 (0=never; 7=several times a day) | | | | | | | | | .01 | .10 |
| Joint activities at home wave 5 (0=never; 7=several times a day) | | | | | | | | | .09 | .09 |
| R² | .01⁺ | | .02 | | .04 | | .10[*] | | .10[*] | |
| adj. R² | .01 | | .01 | | .01 | | .06 | | .06 | |

+p < .10. *p < .05. **p < .01. ***p < .001. Significance of the ANOVA is indicated at R².

which are known to contribute significantly to the prediction of later competencies (LeFevre, Fast et al., 2010; Nguyen et al., 2016).

The analysis of the first research question, to what extent the early home learning environment at the age of two years is associated with mathematical competencies at the age of four years, reveals positive effects of the separate HLE scales on mathematical competencies, irrespective of child and family background variables. Specifically, we find smaller effects of the global stimulation, as indicated in the frequency of joint activities at home, like frequency of shared book reading and singing songs, compared to the domain-specific stimulations, indicated through language and math stimulation during parent-child interactions. However, the significant relations of both parent-child interaction scales and mathematical competencies strengthen the hypothesis of cross-domain effects: Not only math-specific stimulation but also language-specific stimulation are connected to mathematical competencies (Lehl et al., 2020; Napoli & Purpura, 2018). Such cross-domain effects could be explained by the fact that children's mathematical representations are built through language (Wiese, 2003). There is a well-documented relationship between language and mathematics (Peng et al., 2020). In their recent meta-analyses, Peng et al. (2020) conclude that language contributes to math performance via two pathways: through the medium function of language (i.e., using language as a tool for communicating mathematics knowledge with others and building and retrieving representations of mathematics knowledge from long-term memory) and the thinking functions of language (i.e., using language to think about abstract mathematical concepts). Whereas the medium function of language is especially important for foundational mathematics skills (e.g., numerical knowledge and simple calculations), the thinking function of language is particularly important for advanced mathematics. Thus, through high-quality language stimulation at home, parents might foster children's language skills, which in turn might be one pathway to the development of mathematical competencies. Although studies have shown that exposing children to mathematical language during interaction – the so called “math talk” – is highly relevant to mathematical development (Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Ramani et al., 2015), it is still not clear, which language stimulation exactly contributes to mathematics learning.

But the interpretation of the results should take into account that the reliability of the quality indicators of the global HLE ranges from Cronbach's $\alpha = .53$ to $\alpha = .59$ and must be interpreted as quite low. Additionally it has to be mentioned that in wave three, the activity of pretend play was assessed, but this item was not assessed in the further waves. Therefore, the scale of wave three consists of seven items, and the scales of wave four and five consist of six items.

Concerning the question whether later HLE experiences add to or mediate early HLE effects (research question two), we found that the effect of the early joint activities at the child's age of two years vanished when we added the frequency of joint activities at age 3 and 4. As depicted in the research, effects of early years home learning environment are partly mediated by later home learning environment (e.g., Tamis-LeMonda et al., 2019). Our results point in the same direction.

Unfortunately, there was no data on later domain-specific measures of stimulation within the parent-child interaction, yet the global HLE indicator did not mediate the association between the domain-specific indicators and children's mathematical competencies. With regard to the effects of ECEC, we only find a small tendency of the duration of time spent in ECEC, when no control variables are added. The fact that this effect completely vanishes when child and family characteristics are added points to potential selection effects in ECEC. Research shows clear interrelations of parent's educational status or migration background with entry into ECEC, showing that especially parents with a migration background enter institutional child care rather late (Burghardt, 2018). While it was important to add these control variables for exactly this reason, these variables are mainly responsible for the vanishing of the duration-effect.

For the third research question regarding the effects of the frequency of activities in ECEC on children's math outcomes, we only find small effects of extensive activities in institutional child care. These results have to be interpreted with caution. As these tendencies only emerge in the models where control variables are added and were not related to (later) mathematical competencies on their own, the results point to a non-significant effect for frequency of activities in institutional child care. We ran several other statistical analyses as sensitivity analyses, where we included each activity on their own or used a different composite but did not find any significant effects on mathematical competencies. As depicted in the theoretical model, the quality of institutional child care has shown to be of high importance for a child's individual development. The frequency of activities within the whole group of children in institutional child care can only be seen as a rough indicator for the actual process quality, which an individual child experiences during its stay in the childcare-center. Additionally this scale of extensive activities in ECEC shows only small internal validity (Cronbach's $\alpha = .57$). Nonetheless, Linberg et al. (2017) showed that the extensive activities are linked to observed quality.

Furthermore, we do not have any information about whether or not the child was actually involved in the preschool teacher's stated activities nor do we have any information about the quality of the different activities. Additionally, even though the sample started as a nationally representative sample, our analysis sample is quite small; this is particularly true for the third research question regarding effects of the institutional setting. As explained above, this can be traced back to the relatively small sample size of $N = 230$ children, as information on ECEC activities was gained by a drop-off questionnaire for educators, which had a return rate of 31.6 %. Therefore, our results should not be generalized.

Our study contributes to filling research gaps by addressing the early years home and institutional learning environments for children under the age of three in Germany. In summary, we do find effects for global indicators like 'joint activities at home' as well as domain-specific stimulation in 'parent-child interaction'. Following Bronfenbrenner and Morris (2006), these aspects of the home learning environment can be characterized as the "prime engine" in child development. The rather rough indicators for the institutional learning environment calls for further

research. As we could not include any information on stimulation quality or the enrollment of the child in the activities that were assessed in the NEPS, it would be important to conduct a study where process quality could be observed and not assessed in a questionnaire, as observation can be seen as the ‘gold standard’ (Bäumer & Roßbach, 2016). As we also do not know what part mathematical activities play in institutional child care for children under the age of three, even though it is in the educational plans of the federal states (Youth Ministers’ Conference & Culture Ministers’ Conference, 2004), further research should focus on the embedment of math-related processes. As some math intervention programs with preschoolers have been shown to be highly effective (for an overview: Lehrl, 2018), it should be examined which math-specific activities under which circumstances might be profitable for children under the age of three.

References

- Anders, Y., Roßbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*(2), 231–244.
- Aunola, K., Leskinen, E., Lerkkanen, M. K., & Nurmi, J. E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*(4), 699–713.
- Baumeister, A. E., Rindermann, H., & Barnett, W. S. (2014). Crèche attendance and children’s intelligence and behavior development. *Learning and Individual Differences, 30*, 1–10.
- Bäumer, T. & Roßbach, H. G. (2016). Measurement of preschool quality within the National Educational Panel Study: Results of a Methodological Study. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological issues of longitudinal surveys: The example of the National Educational Panel Study* (pp. 543–560). Wiesbaden: VS.
- Blossfeld, H.-P. & Roßbach, H.-G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS)* (2nd ed.). Wiesbaden: VS.
- Blossfeld, H.-P., Roßbach, H.-G., & Maurice, J. von (Eds.). (2011). *Education as a lifelong process. The German National Educational Panel Study (NEPS)*. Wiesbaden: VS.
- Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States part II: Relations with behavioral development through age thirteen. *Child Development, 72*(6), 1868–1886.
- Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic Status and Child Development. *Annual Review of Psychology, 53*, 371–399.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western Society*. New York: Wiley.
- Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 793–828). New York: Wiley.
- Bryant, D. M., Burchinal, M., & Zaslow, M. (2011). Empirical approaches to strengthening the measurement of quality: Issues in the development and use of quality measures in research and applied settings. In M. Zaslow, I. Martinez-Beck, K. Tout,

- & T. Halle (Eds.), *Quality measurement in early childhood settings* (pp. 33–47). Baltimore: Paul H. Brookes.
- Burchinal, M. R., Peisner-Feinberg, E., Pianta, R., & Howes, C. (2002). Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of School Psychology, 40*(5), 415–436.
- Burghardt, L. (2018). Relevance of child characteristics for the envisaged and real start of childcare attendance for the under-3s in Germany. *European Early Childhood Education Research Journal, 26*(6), 873–892.
- Burghardt, L. (2019). Elterliche Einstellungen zum Krippenbesuch und deren Einfluss auf die Inanspruchnahme einer Krippe (Dissertation). Universität Bamberg.
- Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University.
- Clements, D. H. & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for research in Mathematics Education, 136*–163.
- Dornheim, D. (2008). *Prädiktion von Rechenleistung und Rechenschwäche: Der Beitrag von Zahlen – Vorwissen und allgemein-kognitiven Fähigkeiten*. Berlin: Logos.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & Japel, C. (2007). School readiness and later achievement. *Developmental psychology, 43*(6), 1428–1446.
- Eason, S. H. & Ramani, G. B. (2018). Parent–child math talk about fractions during formal learning and guided play activities. *Child Development, 91*(2), 546–562.
- Elliott, L. & Bachman, H. J. (2018). How do parents foster young children’s math skills? *Child Development Perspectives, 12*(1), 16–21.
- Federal Ministry for Family Affairs, Senior Citizens, Women and Youth. (2015). *The ElterngeldPlus with partnership bonus and more flexible parental leave*. Berlin.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314.
- Ganzeboom, H. (2010). *Questions and answers about ISEI-08*. Verfügbar unter: <http://www.harryganzeboom.nl/isco08/qa-isei-08.htm>
- Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: a longitudinal study of process and concept deficits in children with learning disability. *Journal of experimental child psychology, 77*(3), 236–263.
- German Federal Statistical Office. (2019). Kindertagesbetreuung unter Dreijähriger im März 2019: +3,7% gegenüber dem Vorjahr. Verfügbar unter: https://www.destatis.de/DE/Presse/Pressemitteilungen/2019/09/PD19_379_225.html;jsessionid=1FCD9AA06931C4794D534A3AE26EDB6D.internet722
- Grüßing, M., Heinze, A., Duchhardt, C., Ehmke, T., Knopp, E., & Neumann, I. (2013). KiKi – Kieler Kindertagertest Mathematik zur Erfassung mathematischer Kompetenz von vier- bis sechsjährigen Kindern im Vorschulalter. In M. Hasselhorn, A. Heinze, W. Schneider & U. Trautwein (Eds.), *Diagnostik mathematischer Kompetenzen*. (pp. 67–80). Göttingen: Hogrefe.
- Harms, T., Cryer, D., & Clifford, R. M. (2007). *Infant/Toddler Environment Rating Scale – Revised Edition (ITERS-R)*. New York: Teachers College.
- Hauser, B., Vogt, F., Stebler, R., & Rechsteiner, K. (2014). Förderung früher mathematischer Kompetenzen. Spielintegriert oder trainingsbasiert [Promoting early mathematical competence playfully or trainingbased]. *Frühe Bildung, 3*(3), 139–145.
- International Labour Office. (2013). *Working conditions laws report 2012: A global review*. Verfügbar unter: <http://www.ilo.org/dyn/travail>
- Jaffee, S. R., Van Hulle, C., & Rodgers, J. L. (2011). Effects of nonmaternal care in the first 3 years on children’s academic skills and behavioral functioning in childhood and early adolescence: a sibling comparison study. *Child Development, 82*, 1076–1091.

- Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*, 22(1), 36–46.
- Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk". *Developmental Psychology*, 42(1), 59–69.
- Klein, O. & Sonntag, N. (2017). Ethnische Unterschiede der Wirkung institutioneller U3-Kinderbetreuung. *Zeitschrift für Erziehungswissenschaft*, 20(1), 41–60.
- Kluczniok, K., Lehl, S., Kuger, S., & Roßbach, H.-G. (2013). Quality of the home learning environment during preschool age – Domains and contextual conditions. *European Early Childhood Education Research Journal*, 21(3), 420–438.
- Knopp, E., Duchhardt, C., Ehmke, T., Grüßing, M., Heinze, A., & Neumann, I. (2014). Von Mengen, Zahlen und Operationen bis hin zu Daten und Zufall – Erprobung eines Itempools zum Erfassen der mathematischen Kompetenz von Kindergartenkindern. *Zeitschrift für Grundschulforschung*, 7(1), 20–34.
- Krajewski, K., Grüßing, M., & Peter-Koop, A. (2009). Die Entwicklung mathematischer Kompetenzen bis zum Beginn der Grundschulzeit. In A. Heinze & M. Grüßing (Eds.), *Mathematiklernen vom Kindergarten bis zum Studium: Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (p. 17–34). Münster: Waxmann.
- LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, 81, 1753–1767.
- LeFevre, J.-A., Polyzoi, E., Skwarchuk, S.-L., Fast, L., & Sowinski, C. (2010). Do home numeracy and literacy practices of Greek and Canadian parents predict the numeracy skills of kindergarten children? *International Journal of Early Years Education*, 18(1), 55–70.
- Lehl, S. (2018). *Qualität häuslicher Lernumwelten im Vorschulalter. Eine empirische Analyse zu Konzept, Bedingungen und Bedeutung*. Wiesbaden: VS.
- Lehl, S., Ebert, S., Blaurock, S., Roßbach, H.-G., & Weinert, S. (2020). Long-term and domain-specific relations between the early years home learning environment and students' academic outcomes in secondary school. *School Effectiveness and School Improvement*, 31(1), 102–124.
- Lehl, S., Kluczniok, K. & Roßbach, H.-G. (2016). Longer-term associations of preschool education: The predictive role of preschool quality for the development of mathematical skills through elementary school. *Early Childhood Research Quarterly*, 36, 475–488.
- Lehl, S., Kluczniok, K., Roßbach, H.-G., & Anders, Y. (2017). Long term persistence of preschool intervention on children's mathematical development: Results from the German model project "Kindergarten of the future in Bavaria". *Global Education Review*, 4(3), 70–87.
- Leseman, P. P. M. & de Jong, P. F. (1998). Home literacy: Opportunity, instruction, cooperation and social-emotional quality predicting early reading achievement. *Reading Research Quarterly*, 33(3), 294–318.
- Linberg, A. (2018). *Interaktionen zwischen Mutter und Kind. Dimensionen, Bedingungen und Effekte*. Empirische Erziehungswissenschaft. Münster: Waxmann.
- Linberg, A., Kluczniok, K., Burghardt, L., & Freund, J.-D. (2017). Quality of toddler childcare – Can it be assessed with questionnaires? *Early Child Development and Care*, 189(8) 1369–1383.
- Linberg, A., Lehl, S., & Weinert, S. (2020). The early years home learning environment – Associations with parent-child-course attendance and children's vocabulary at age 3. *Frontiers in Psychology*, 11, 1425.

- Linberg, A., Mann, D., Attig, M., Vogel, F., Weinert, S., & Roßbach, H.-G. (2019). *Assessment of interactions with the macro-analytic rating system of parent-child-interactions in the NEPS at the child's age of 7, 17, and 26 months* (NEPS Survey Paper No. 51). Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Loeb, S., Fuller, B., Kagan, S. L., & Carrol, B. (2004). Childcare in poor communities: Early learning effects of type, quality, and stability. *Child Development*, 75(1), 47–65.
- Melhuish, E. C., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., & Leseman, P. (2015). A review of research on the effects of early childhood education and care (ECEC) on children development. *CARE Curriculum and quality analysis impact review*, WP4.1.
- Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues*, 64(1), 95–114.
- Mistry, R. S., Benner, A. D., Biesanz, J. C., Clark, S. L., & Howes, C. (2010). Family and social risk, and parental investments during the early childhood years as predictors of low-income children's school readiness outcomes. *Early Childhood Research Quarterly*, 25(4), 432–449.
- Napoli, A. R. & Purpura, D. J. (2018). The home literacy and numeracy environment in preschool. Cross-domain relations of parent-child practices and child outcomes. *Journal of Experimental Child Psychology*, 166, 581–603.
- Neumann, I., Duchardt, C., Grüßing, M., Heinze, A., Knopp, E., & Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for Educational Research Online*, 5(2), 80–109.
- NICHD Early Child Care Research Network. (1999). *Mother-child-interaction – semi-structured procedure (six-month home visit): Phase 1 instrument document*. Unpublished manuscript.
- NICHD Early Child Care Research Network. (2002a). Child-care structure – process – outcome: Direct and indirect effects of childcare quality on young children's development. *Psychological Science*, 13(3), 199–206.
- NICHD Early Child Care Research Network. (2002b). Early childcare and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Education Research Journal*, 39, 133–164.
- NICHD Early Child Care Research Network. (2006). Child-care effect sizes for the NICHD Study of Early Child Care and Youth Development. *American Psychologist*, 61(2), 99–116.
- NICHD Early Child Care Research Network & Duncan, G. J. (2003). Modeling the impacts of childcare quality on children's preschool cognitive development. *Child Development*, 74(5), 1454–1475.
- Niklas, F. & Schneider, W. (2012). Einfluss von Home Numeracy Environment auf die mathematische Kompetenzentwicklung vom Vorschulalter bis Ende des 1. Schuljahres. *Zeitschrift für Familienforschung*, 24, 25–41.
- Nguyen, T., Watts, T. W., Duncan, G. J., Clements, D. H., Sarama, J. S., Wolfe, C., & Spitler, M. E. (2016). Which preschool mathematics competencies are most predictive of fifth grade achievement? *Early childhood research quarterly*, 36, 550–560.
- OECD – Organisation für wirtschaftliche Zusammenarbeit und Entwicklung. (2013). *How do Early Childhood Education and Care (ECEC) policies, systems and quality vary across OECD countries?* Paris: OECD.
- Peng, P., Lin, X., Ünal, Z. E., Lee, K., Namkung, J., Chow, J., & Sales, A. (2020). Examining the mutual relations between language and mathematics: A meta-analysis. *Psychological Bulletin*, 146(7), 595–634.

- Ramani, G. B., Rowe, M. L., Eason, S. H., & Leech, K. A. (2015). Math talk during informal learning activities in Head Start families. *Cognitive Development, 35*, 15–33.
- Rodriguez, E. T., Tamis-LeMonda, C. S., Spellmann, M. E., Pan, B. A., Raikes, H., Lugo-Gil, J., & Luze, G. (2009). The formative role of home literacy experiences across the first three years of life in children from low-income families. *Journal of Applied Developmental Psychology, 30*(6), 677–694.
- Sammons, P., Toth, K., Sylva, K., Melhuish, E., Siraj, I., & Taggart, B. (2015). The long-term role of the home learning environment in shaping students' academic attainment in secondary school. *Journal of Children's Services, 10*(3), 189–201.
- Sénéchal, M. & LeFevre, J.-A. (2002). Parental Involvement in the development of children's reading skill: a five-year longitudinal study. *Child Development, 73*(2), 445–460.
- Skillen, J., Berner, V.-D. & Seitz-Stein, K. (2018). The rule counts! Acquisition of mathematical competencies with a number board game. *The Journal of Educational Research, 111*(5), 554–563.
- Spelke, E. S. (2017). Core knowledge, language, and number. *Language Learning and Development, 13*(2), 147–170.
- Susperreguy, M. I., Di Lonardo Burr, S., Xu, C., Douglas, H., & LeFevre, J. A. (2020). Children's Home Numeracy Environment Predicts Growth of their Early Mathematical Skills in Kindergarten. *Child Development 91*(5), 1663–1680.
- Sylva, K., Melhuish, E., Sammons, P., Siraj, I., & Taggart, B. (2014). *Effective preschool, primary and secondary education 3–16 Project (EPPSE 3–16) – Students' educational and developmental outcomes at age 16*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/351496/RR354_-_Students__educational_and_developmental_outcomes_at_age_16.pdf
- Sylva, K., Stein, A., Leach, P., Barnes, J., & Malmberg, L.-E. (2011). Effects of early child-care on cognition, language, and task-related behaviours at 18 months: an English study. *British Journal of Developmental Psychology, 29*, 18–45.
- Tamis-LeMonda, C. S., Luo, R., McFadden, K. E., Bandel, E. T., & Valloton, C. (2019). Early home learning environment predicts children's 5th grade academic skills. *Applied Developmental Science, 23*(2), 153–169.
- Tietze, W., Becker-Stoll, F., Bense, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B., & Leyendecker, B. (2013). *Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK)*. Weimar: Verlag das Netz.
- Ulferts, H. & Anders, Y. (2016). *Effects of ECEC on academic outcomes in literacy and mathematics: Meta-analysis of European longitudinal studies*. Retrieved from http://ecccare.org/fileadmin/careproject/Publications/reports/CARE_WP4_D4_2_Metaanalysis_public.pdf
- Youth Ministers' Conference & Culture Ministers' Conference (2004). *Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen* (Beschluss der Jugendministerkonferenz vom 13./14.05.2004 / Beschluss der Kultusministerkonferenz vom 03./04.06.2004). Retrieved from https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_06_03-Fruhe-Bildung-Kindertageseinrichtungen.pdf
- Waldfogel, J. (2001). International policies toward parental leave and child care. *Caring for Infants and Toddlers, 11*, 99–111.
- Wiese, H. (2003). *Numbers, language, and the human mind*. Cambridge: Cambridge University.
- Wood, D., Wood, H., & Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. *International Journal of Behavioral Development, 1*(2), 131–147.
- Wygotski, L. & Cole, M. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University.

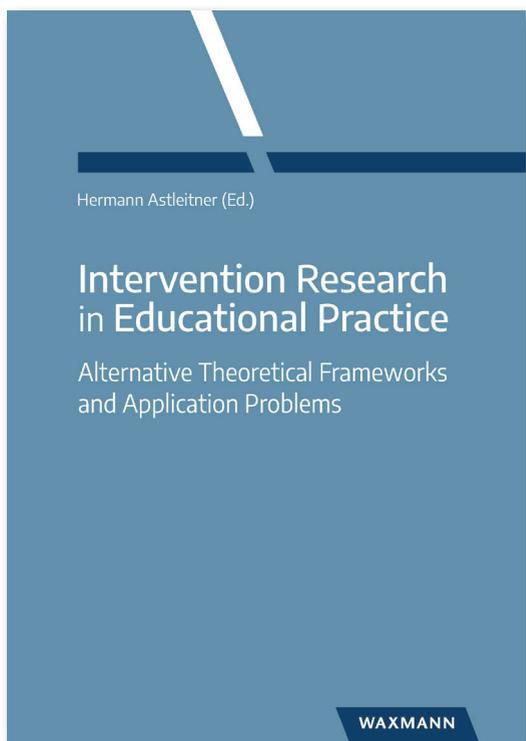
- Wylie, C., Ferral, H., Hodgen, E., Thompson, J., & New Zealand Council for Educational Research. (2006). *Contributions of early childhood education to age-14 performance. Evidence from the competent children, competent learners project*. Wellington: New Zealand Council for Educational Research.
- Zippert, E. L. & Rittle-Johnson, B. (2020). The home math environment: More than numeracy. *Early Childhood Research Quarterly*, 50, 4–15.

Hermann Astleitner (Ed.)

Intervention Research in Educational Practice

Alternative Theoretical Frameworks and Application Problems

2020, 188 pages, pb, € 29,90,
ISBN 978-3-8309-4197-2
E-Book: € 26,99
ISBN 978-3-8309-9197-7



With contributions by

Hermann Astleitner, Josef Eisner,
Ulrike Greiner, Franz Hofmann,
Michaela Katstaller, Nomy Dickman,
Gabriela Gniewosz, Stephanie
Musow, Alexander Naumann

Improving learning and teaching at schools or universities may start with choosing evidence-based interventions and practices, but does not end there. To ensure sustainable changes to programs in educational practice, interventions need to address complex issues related to theories, research designs, and measurements. This book presents typical but often overlooked problems in intervention research in educational practice. These problems are embedded in various educational areas such as, amongst others, school effectiveness, instructional design or motivational aspects of teacher trainings.

WAXMANN

www.waxmann.com
info@waxmann.com

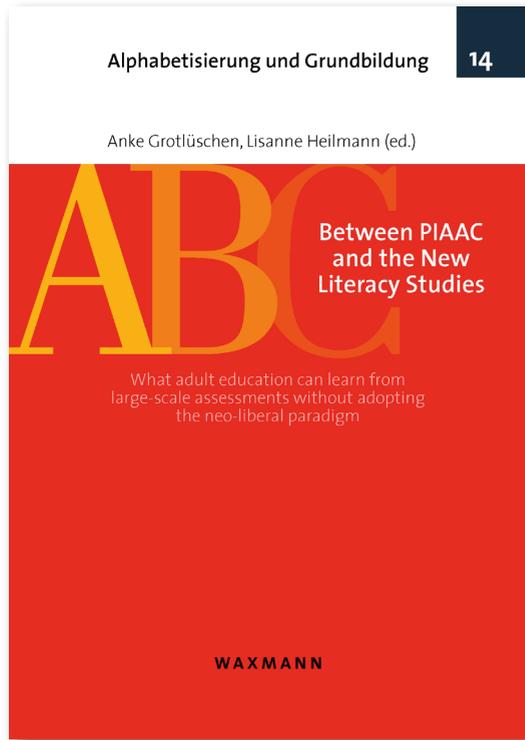
Anke Grotlüschen, Lisanne Heilmann (ed.)

Anke Grotlüschen,
Lisanne Heilmann
(Eds.)

Between PIAAC and the New Literacy Studies

What adult education can learn
from large-scale assessments
without adopting the
neo-liberal paradigm

Alphabetisierung und Grundbildung,
vol. 14, 2020, ca. 270 pages, pb.,
€ 34,90, ISBN 978-3-8309-4188-0
E-Book: open access,
ISBN 978-3-8309-9188-5
doi.org/10.31244/9783830991885



With contributions by

Klaus Buddeberg,
Svetlana Chachashvili-Bolotin,
Caroline Duncker-Euringer,
Gregor Dutz, Anke Grotlüschen,
Lisanne Heilmann,
Barbara Nienkemper, Thomas J. Sork
Christopher Stammer, Jana Wienberg

With this book we present a selection of articles that critically deal with (internationally comparative) large-scale assessments. We acknowledge that studies such as PIAAC are often designed, financed and implemented on the basis of neo-liberal worldviews. Nevertheless, we would like to use the articles that are presented here to show the various ways in which adult and continuing education can benefit and learn from the knowledge that they generate. In PIAAC, for example, there are huge differences between the surveyed variables and the theoretical frameworks on literacies and literacy practices that the New Literacy Studies (NLS) have brought out. This book features eleven articles, which – with the NLS's theoretical considerations and points of criticism in mind – find new and alternative evaluations and interpretations of the data. Not only can they show effects of marginalization on a large scale, but the data can also provide information about mechanisms of power in relation to literacy and basic competencies.

WAXMANN

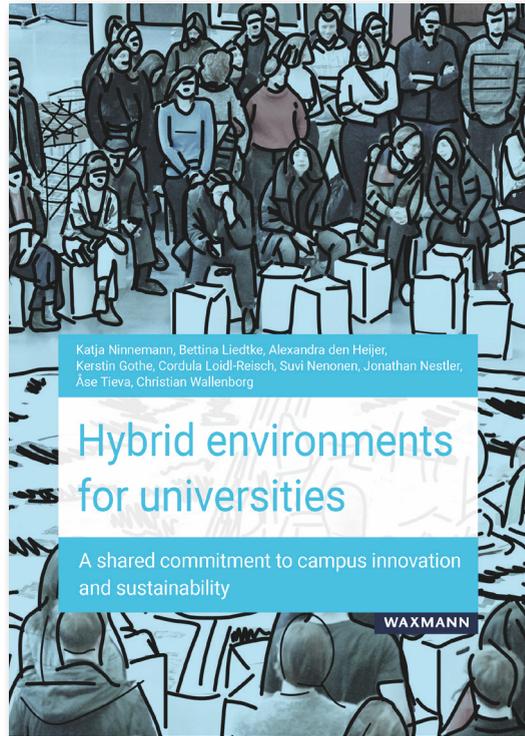
www.waxmann.com
info@waxmann.com

Katja Ninnemann, Bettina Liedtke,
Alexandra den Heijer, Kerstin Gothe,
Cordula Loidl-Reisch, Suvi Nenonen,
Jonathan Nestler, Åse Tieva,
Christian Wallenborg

Hybrid environments for universities

A shared commitment
to campus innovation
and sustainability

2020, 120 pages, pb., € 19,90,
ISBN 978-3-8309-4179-8
E-Book: Open Access
doi.org/10.31244/9783830991793



This publication is the result of an international and interdisciplinary expert meeting at Technische Universität Berlin, in March 2020. The aim of the expert meeting was to collaboratively write and publish a book, within five days, on the central question: Which organizational structures and processes at universities support a strategic as well as innovative campus development?

As experts with an interdisciplinary background including the social sciences, public real estate, urban planning, architecture and landscape architecture, we could examine the question from a holistic perspective and gain new insights.

The resulting manifesto states necessary steps and strategies to create innovative and sustainable hybrid environments for universities. It addresses all decision makers – executives, practitioners and contributors alike – as all of us face the challenge of limited resources and needing to do more with less.

WAXMANN

www.waxmann.com
info@waxmann.com