Program Implementation and Effectiveness of Extracurricular Activities: An Investigation of Different Student Perceptions in Two German All-Day Schools

Stephan Kielblock

Abstract: During the past decade, many schools in Germany have added extracurricular time to their regular curricular classes. This raises questions about the successful implementation of extracurricular programs and what makes them effective. The aim of this study is to illuminate the connection between these two questions. The theoretical and conceptual framework suggests that individual perception is a core concept that links both issues. Based on multi-method data from the Study on the Development of All-day Schools (StEG), the individual perceptions of two different activities will be investigated. One activity is perceived by the students as just an "extra" curricular activity (an extension of regular classes), whereas the other activity is seen as a real new "extracurricular" opportunity. The results emphasise the importance of viewing student perceptions in a qualitative manner.

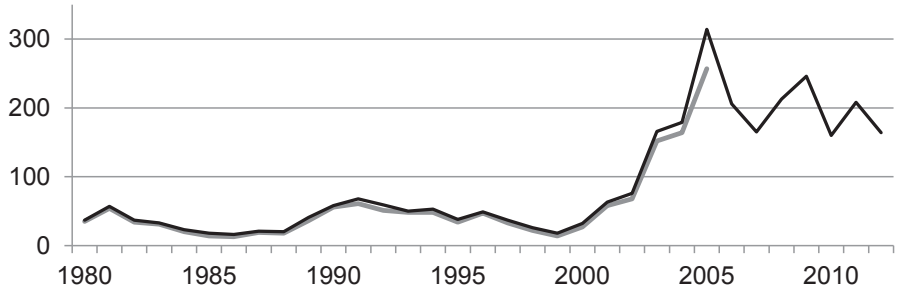
Keywords: program implementation, effectiveness, student perception, all-day school, extracurricular activities

Introduction and research question

Research efforts in the field of German all-day schools[1] have intensified over the past decade in Germany. This can be determined by a literature search using the data base of FIS-Bildung. Papers, articles and books tagged with the keyword *Ganztagsschule* (all-day school) are presented in figure 1. Similar to Holtappels et al. (2008), figure 1 shows that the annual amount of literature dealing with all-day schools has increased in particular since 2003. In recent years, the number of such publications has levelled off, yet remains relatively high.

1 A short summary in English on how an all-day school is defined in Germany can be found e. g. in Fischer & Klieme (2013, especially p. 29).

Figure 1. Annual number of publications concerned with all-day schools (Ganztagsschule) in Germany from 1980 to 2012. Source: Literature search using the FIS-Bildung database; annual number of publications tagged with the keyword: Ganztagsschule (all-day school). Literature search on 18 February 2014 (dark graph line) and Holtappels et al. (2008, p. 37; light graph line) in comparison.



In Germany, the PISA 2000 study resulted in an intensive public and academic dialogue about the appropriateness of the German educational system. Since 2003, massive public funds (especially the IZBB program; 2003–2009) have been granted to change former half-day schools into all-day schools and to build new all-day schools. The result has been an almost linear increase in the number of all-day schools (Kielblock & Stecher, 2014). More than half of all schools (or more precisely "administrative divisions") are now organised as all-day schools in Germany. The most recent statistical data (KMK, 2006–2014) shows an average of 1,080 new all-day schools in Germany per year for the past decade. The annual increase of new all-day schools fluctuated between almost 400 (between 2011 and 2012) to up to more than 1,500 (between 2008 and 2009).

To evaluate this expansion, the Study on the Development of All-day Schools (StEG) was established in 2005. At the end of the first StEG project phase, which lasted from 2005 to 2011, Klieme and Rauschenbach (2011) concluded that the elaborated multi-perspective and longitudinal design of the StEG produced differentiated knowledge. Yet, more specific knowledge is still necessary in order to provide profound recommendations for educational policy on how to implement attractive, high-quality extracurricular activities (Klieme & Rauschenbach, 2011, p. 349). Although it is predictable that growth in the number of new all-day schools per year will eventually decrease, two major issues in Germany's all-day schools continue to be (1) how to successfully implement extracurricular activities in schools and (2) what makes them effective.

Both of these are very difficult questions that cannot reasonably be answered in a single paper. However, according to Klieme and Rauschenbach (2011), the question of implementation and the question of effectiveness appear to be connected. Consequently, the aim of this paper is to illuminate whether and how these two questions have a common basis. Therefore, the research question is how the problem of successful program implementation is related to the issue of program effectiveness.

Program implementation: a literature search

Successful program implementation is a major topic of implementation science. Kelly states that "implementation science is the study of the processes and methods involved in the systematic transfer and uptake of evidence-based practices into routine, everyday practice" (Kelly, 2012, p. 4). Although it is a relatively new approach to investigate the implementation of innovations in real world contexts (Kelly, 2012, p. 3), there are at least some common core characteristics and ideas. Implementation is seen as a process that can be systematised into different stages. The progress through these different stages is promoted by competency drivers (e. g. coaching), organisation drivers (e. g. decision-support data system) and sufficient leadership (Blase et al., 2012, p. 16).

Since research on the implementation of extracurricular activities or programs in German educational literature is not very extensive, an intensive literature search was performed to consider international literature in identifying relevant papers. The Education Resources Information Center (ERIC) indexes a wide variety of international journal sources, so this database was searched using the term "implementation." The following criteria led to a significant number of published papers concerning program implementation from the past ten years. Three main issues were considered in the selection of sources. First, these papers should focus on the implementation of a specific non-curricular program, activity or intervention. Second, papers are included that have a focus on schools in a broader sense. For example, papers dealing with higher education or early childhood have been excluded. And finally, the programs described in the papers should primarily target students.

The literature reveals both facilitators and barriers of program implementation. Most importantly, the implementation of a new program must fit into an established system or curriculum. Therefore, new ideas could possibly collide with existing standards (Greaney, et al., 2007, p. 254) and could present a significant change from traditional patterns (Schwalm & Tylek, 2012). For example, Olvera et al. (2008) found that one barrier for program implementation was mandatory school tutorials during afterschool time, which hindered female students in attending the new BOUNCE program. Similar conflicts with already existing afterschool programs, students' competing priorities or an unalterable infrastructure (e. g. school busses) are also reported in other studies (Schwalm & Tylek, 2012; Greaney et al., 2007). In these cases, problems arose because the new program upset a well-established daily routine. That "each school's rhythm must be respected" (Deslandes, 2006, p. 102) is a wise notion for program implementation in this respect, yet it is not easy to follow. One idea for solving the problem of fitting in with the school's rhythm is to integrate the new program into the school's developmental (or improvement) plan so that it becomes a priority (Deslandes, 2006, p. 100). Choosing program content appropriate for the needs of students and the community (Grimmett, Rickard, & Gill, 2010, p. 61) and keeping an eye on the local context (McIntyre et al., 2005, p. 89) are other ideas that can be found in the literature. The external environment or the community is also discussed in several studies that reflect program implementation against the

backdrop of the school context (Huang et al., 2009; Olvera et al., 2008; Grimmett, Rickard, & Gill, 2010).

The literature search also reveals the facilitators and barriers of program implementation with respect to the importance of leadership and effective management (Deslandes, 2006; Grimmett, Rickard, & Gill, 2010; Hall, 2010). In a wider sense, the director of the afterschool program is seen as the backbone due to the motivation of teachers to support program implementation (Greaney et al., 2007; McIntyre et al., 2005). In concurrence, Huang et al. (2009) report in their analyses of a health promotion program in Taiwan "that the major guiding force was from the school principal and head of academic affairs office" (Huang et al., 2009, p. 94).

Yet, the principal is not the only person who encourages all of the stakeholders to support program implementation. Deslandes (2006) speaks of liaison agents "with a stable and credible relationship with other players in the school and with demonstrated motivational skills" (Deslandes, 2006, p. 101). In addition, other stakeholders such as students, teachers, other staff, parents and other community members are seen as crucial for program implementation (Grimmett, Rickard, & Gill, 2010; Greaney, et al., 2007; Davis & Clark, 2012; Hallenbeck & Fleming, 2011; Schwalm & Tylek, 2012). All of them must understand the program's purposes (Huang et al., 2009), share a common vision and work cooperatively (Ocak, 2011). This coincides with the findings of Collier and Henriksen (2012) that "success for program implementation by teachers depends, in part, upon how comfortable they are with the approach and how motivated they are about the approach" (Collier & Henriksen, 2012, p. 14).

One further aspect concerning the facilitators and barriers of program implementation emerged in the literature search. Deslandes (2006) derives from her empirical analyses in Canada that patience is very important for successful implementation. There must be "time to become familiar with the project" (Deslandes, 2006, p. 101). In their data from the US, Greaney et al. (2007) find time constraints that also challenged successful implementation. Yet, the literature not only reports time as a scarce resource, but also funds and money (Deslandes, 2006, p. 101; Greaney et al., 2007, p. 255; Grimmett, Rickard, & Gill, 2010, p. 61). As Ocak (2011) summarises: "A considerable investment of time, effort, resources, and money" (Ocak, 2011, p. 1399) is needed.

Additional evaluation of organisational achievements, as reflected in Huang et al. (2009, p. 95), should also be seen as a facilitating component of program implementation. By and large, evaluation should help to understand the results of specific program implementation processes and should therefore aid in planning further adjustments of the implemented program or inform about the implementation of other programs. However, Zhang et al. (2011) criticise the common summative evaluation practice and suggest that the best evaluation approach should "systematically guide both evaluators and stakeholders in posing relevant questions and conducting assessments at the beginning of a project […], while it is in progress […], and at its end" (Zhang et al., 2011, p. 59). Consequently, Zhang et al. (2011, p. 61) propose using Stufflebeam's Context, Input, Process, and Product evaluation model (CIPP). On the one hand, there is the question as to what the different stakeholders need (context evaluation) and how these needs are addressed through the program (input evalua-

tion) (Zhang et al., 2011). On the other hand, product evaluation "measures, interprets, and judges project outcomes and interprets their merit, worth, significance, and probity" (Zhang et al., 2011, p. 59). Between the context and input on the one side and product on the other, process evaluation "monitors the project process and potential procedural barriers, and identifies needs for project adjustments" (Zhang et al., 2011, p. 59).

The main problem with process evaluation is to not misinterpret it as an earlier or repeated product evaluation. The literature search presented here includes neither solid nor reliable examples of a process evaluation. This could be due to resource constraints, such as the fact that formative assessment takes exorbitantly more time, effort, resources and money in comparison to summative assessment, which only illuminates the products of the processes. This is true for all stakeholders involved in the ongoing processes. One illustration of the investigation of these processes can be found in Zhang et al. (2011, p. 73–74). Biweekly meetings, observations and several curriculum-based measure probes were introduced to monitor service learning program implementation. Although Zhang et al. (2011) do not explicitly discuss the amount of work, it can be assumed from their descriptions that process-accompanying evaluation could be quite exhausting for all stakeholders.

In addition to this practical consideration of process evaluation, methodological issues arise. The evaluation of specific *states* of ongoing *processes* (for example, several curriculum-based measure probes; Zhang et al., 2011) establishes the imperative that an evaluator must have a good theoretical representation of what happens during the processes of implementation. Otherwise, process evaluation is mistakenly informed solely by program goals and therefore underestimates the complexity of ongoing change processes.

An interesting insight into change and implementation processes is given by Hall (2010) with his metaphor of the "implementation bridge." This metaphor explains that "current practices in schools and classrooms" and "new practices" are divided by a chasm and that only an "implementation bridge" enables the movement from current to new practices. This bridge consists of three layers that illustrate the different processes on their way from one side to the other. These layers are "stages of concern" (stages from "unconcerned" up to "refocusing"), "levels of use" (stages from "non-use" up to "renewal") and "innovation configurations" (stages from "no fidelity" up to true implementation "fidelity") (Hall, 2010). On the whole, "the researcher/evaluator can use information from each dimension to measure how far across the bridge each implementer has progressed. Change facilitators can also use the same constructs and information for planning and making interventions to help implementers move further across the bridge" (Hall, 2010, p. 235).

The components of this metaphor seem very reasonable. It also seems to be true that "the extent and quality of use for new approaches can be greatly enhanced when there is understanding of how people change" (Hall, 2010, p. 232). Yet, in consideration of the notion that it is important to understand how people change and how the complexity of implementation processes can be theorised, a genuinely process-oriented perspective on students seems to be omitted by Hall (2010, as well as by others, e. g. Skaggs & Bodenhorn, 2006; Everhart, 2005). The literature search presented here highlights this gap. Therefore, the next section attempts to bridge this

gap by introducing a model for illuminating student processes that lead to student outcomes.

Processes leading to student outcomes

In the previous section, the literature review suggests that the students' perspective is commonly represented not as a process but only as a result of other processes that lead to student outcomes. Therefore, it seems that student-specific processes leading to student outcomes are barely considered by literature on program implementation. For example, this is true for Hall (2010), who has an elaborated multi-layer approach to implementation with a specific focus on teacher's views and teacher actions that impact students' outcomes. Hall (2010) calls on researchers to investigate student outcomes ("test scores") with regard to "how far across the [implementation] bridge each implementer has moved" (Hall, 2010, p. 251). However, if more or less successfully implemented "new practices" lead to student outcomes (or not), it remains unclear what exactly happened to the students. Consequently, it is assumed here that the reflection on program implementation might benefit from understanding the student-specific processes leading to specific student outcomes. As a result, this section focusses on introducing a model that genuinely considers student outcomes as a process.

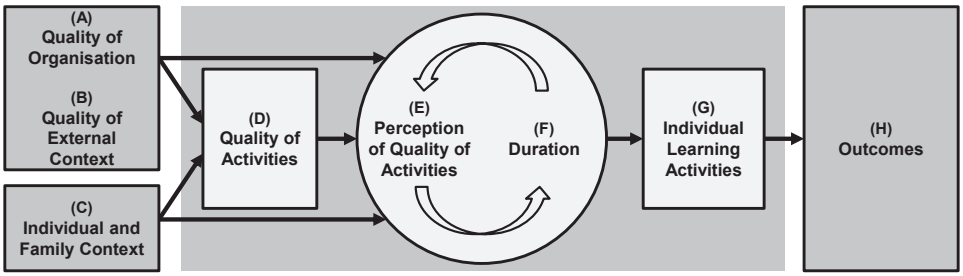
A model commonly used by German researchers concerned with the effectiveness of extracurricular activities at all-day schools is the "model of educational quality of extracurricular activities in all-day schools" (*Modell der Bildungsqualität ausserunterrichtlicher Angebote in der Ganztagsschule*) that Stecher et al. (2007, p. 350) derived mainly from the reflections of Miller (2003, p. 43) on US afterschool programs. After years of the model being refined in the German-speaking community (e. g. Radisch et al., 2008; Radisch, 2009; or Fischer et al., 2012), it actually re-entered international discourse (e. g. in Fischer & Klieme, 2013 or in Stecher & Maschke, 2013).

The basis of the model is an input-process-outcome relation. The input or context perspective contains the quality of the school, the external context (e. g. cooperation partners of the school) and the individual and family context. The process perspective focuses on two different aspects of extracurricular activities. On the one hand, the process characteristics of the activities are modelled by means of the basic dimensions of pedagogical process quality. On the other hand, they consist of measuring attendance for the extracurricular activities in terms such as the absolute attendance, intensity, duration and breadth of attendance. The input or context perspective and the process perspective result in the outcome perspective that includes educational outcomes, school achievement and school attainment.

Another model commonly used by German researchers concerned with the effectiveness of pedagogical contexts is the "offer and use model" (Helmke, 2003, 2009). This model breaks with the tradition of having direct paths from instruction to outcomes; instead, individual perception and interpretation is crucial for individual action. Placing an emphasis on the ideas of Helmke (2003, 2009), the model of

educational quality (Stecher, et al., 2007) requires specific revisions that are described in the following paragraphs. Figure 2 shows the merging of the two models. The dark grey colour of the model components indicates the input and outcome perspective, whereas the light grey colour of the model components indicates the processes that lead from input to outcome.

Figure 2. Generalised model of educational quality concerning activities in the field of extended education. Own development based on the models of Stecher et al. (2007) and Helmke (2009)



On the one hand, the model depicts that the activity (or activities) of interest is (are) implemented or embedded in a specific organisational quality. Therefore, the "quality of organisation" (A) is one of the starting points. The "external context" (B) of the organisation provides resources that are immediately connected to the opportunities of the organisation. This is why both of them are joined together in just one model component. On the other hand, the "individual and family contexts" (C) of the participants affect the ongoing processes as well. So they are also part of the input perspective. When taken together, these model components (A, B & C) actually have an impact on the "quality of activities" (D). They also affect how individuals "perceive"[2] (E) the quality of activities, as well as affecting the intensity, duration and breadth of attendance (F). Out of the recursive process between E and F, the student delves into individual learning activities (G) that facilitate specific outcomes (H). As Helmke (2003, 2009) emphasises, the individual student must become actively involved in the learning processes. Otherwise, significant student outcomes are not likely.

The processes shown in figure 2 reflect the basic ideas of Stecher et al. (2007). Yet, two new model components are introduced: the perception of the quality of the activities (E) and the individual learning activities (G). The former is an extension of the "quality of activities" (D) connected with the three "basic dimensions of instructional quality" (Klieme, et al., 2006; Kunter, et al., 2007; and in English e. g. Lipowsky, et al., 2009) that include cognitive activation, supportive climate and classroom management. The common conception of the "quality of activities" (D) is a standardised measure (the dimensions are predefined) and a global measure (the standardised indicators target students in general or the "activity" but not how the individual feels during the activities or how he/she individually defines his/her situation "to be in the activity" and which aspects of the quality are relevant to the in-

2 The German term *Wahrnehmung* is difficult to translate literally because it corresponds very much to the English term "perception" yet is also related to meanings of "experience."

dividual). This is the central focus of the offer-and-use model (Helmke, 2003, 2009). It doubts that the quality of activities could have an *immediate* effect on students. Therefore, Helmke (2003, 2009) separates classes ("offer") and learning activities ("use") by a mediating model component, namely the individual perception and interpretation (German: *Wahrnehmung und Interpretation*). In this sense, it is not so important to the individual how he/she describes the activities (What is the activity like?). Yet, how he/she reflects his/her own stance within the activities (What is my position within the activity? What is my situation in the activities and is it a significant experience with regard to my educational biography?) is crucial (this differentiation is also discussed e. g. in Klerfelt & Haglund, 2014).

From this point of view, it becomes clear why attendance and its intensity, duration and breadth (see Fiester, Simpkins & Bouffard, 2005) are closely and recursively connected to the "perception of the quality of activities." The individual grows into the setting of the specific activity; in becoming more and more socialised within that setting, the perception of the quality of activities becomes more affected. This is why the "perception of the quality of activities" (E) is modelled in a recursive process with the "duration" (F) in figure 2. And because of this the term 'duration' is used in the model for the different aspects of attendance.

Theoretical backing for the model is provided by methodological individualism (see e. g. Coleman, 1990). The transition in figure 2 from D to E/F corresponds to what is called the "logic of the situation." This means that the individual perceives or defines a situation (for example, in a certain activity) to be "new" or at least as "odd" in comparison to the daily routine. In defining a situation as "new," the individual faces an altered set of possible actions. One chosen path of action corresponds to the transition from E/F to G in figure 2. In terms of methodological individualism, the chosen action and the action itself are called the "logic of selection." The aggregate of all actions is called the "logic of aggregation" and is represented by the transition from G to H.

As the literature review revealed, the processes leading to specific student outcomes through program implementation are mainly illuminated by observing the pedagogical staff and their individual paths towards implementation. This section focuses on and emphasises the student's point of view. From this perspective, the conviction emerges that particularly the recursive process between the individual perception of the quality of activities and their duration should be considered as important. Consequently, the proposition arising from this is that actual gains in student outcomes may possibly become achievable through individual learning activities only if students see their situation in the program of interest as new and significant. Conversely, if a program is interpreted by students to be a daily routine, it is very unlikely for considerable outputs to be produced that could possibly be traced back to the program of interest.

According to the research question formulated at the end of section one, how the problem of a successful program implementation is related to the problem of their effectiveness holds considerable interest. When informed by the theoretical and conceptual framework presented in sections two and three, it becomes evident that change processes on an "institutional level" (such as the implementation of a

program) are based on the processes of perception, definition and interpretation (and action) by students on an "individual level."

To support these theoretical and conceptual considerations, it is necessary to have empirical evidence as to how students *perceive* different activities. To be more precise with regards to figure 2, the transformation from D to E/F and from E/F to G holds specific interest. The empirical data presented in the following might give initial clues in this direction as to how a specific "offer" (D) is transformed into subjective perceptions (E/F) and how these lead to action (G).

Methods

As mentioned in the introduction, the first project phase of the StEG (2005-2011) was very productive in generating broad knowledge about all-day schools in Germany. This is why the current project phase of the StEG (2012-2015) concentrates on more specific issues and aims for a much closer focus on the effectiveness of extracurricular time. One of the sub-projects of the StEG (called StEG-Q) concentrates on case studies of specific extracurricular activities.

The StEG-Q research team investigated various extracurricular activities at all-day schools in Hesse, a German state. Two years of field work were planned for 2013 and until the end of 2014. For this paper, two extracurricular activities were chosen at two different all-day schools: the first is a Cooperative Comprehensive School and the second is a German *Gymnasium*[3]. Both schools may differ with respect to organisational structure, but they have comparable elements with regard to their efforts at integrating what has traditionally been considered "homework" into the extracurricular time of the school day. Because of the extended school day at all-day schools, homework has become no longer feasible. Therefore, both schools have implemented special extracurricular activities to replace traditional homework done in the afternoon at home.

In addition to interviews with those conducting the activities, 36 students at the two selected schools also participated in the StEG-Q study. There were 19 girls and 17 boys. Most of them were from 10 to 11 years old. Only one student was 9 and two students were 12. Because no quantitative data was collected on all students of the participating schools, it can only be speculated that the *Gymnasium* may have a different composition in comparison to the Cooperative Comprehensive School regarding students' SES, for example. But what is commonly referred to as "selection bias" in qualitative inquiries leads here to two quite similar groups of students who individually agreed and whose parents also agreed to participate in the long-term StEG-Q investigation. So there are no considerable differences between the participating students with regard to their number, gender or age.

3 A *Gymnasium* is one form of secondary education in Germany, the completion of which enables students to further pursue an academic (university) degree. This is not necessarily the case in cooperative comprehensive schools as three different qualifications are possible that enable further education at either trade, professional or academic institutions.

Data was collected in the autumn of 2013. At both schools, multiple methods were utilised to investigate extracurricular activities. First, data was collected by interviewing the member of the pedagogic personnel providing the activity in question. In a semi-structured interview format, the staff members were asked by trained interviewers to talk about their intents and ideas of their extracurricular activity. For the most part, interviewers followed the instructions of problem-centred interviews (Witzel & Reiter, 2012; Kielblock & Lange, 2013) as a brief guide that roughly informs the progress of the conversation. Second, group discussions with students attending those extracurricular activities were carried out in order to investigate the collective student perspective. The methodological discourse brought the "focus groups" (Barbour, 2007) more into line with German discourse on "group discussions." Yet, there were at least some differences (Bohnsack, 2004). Instructions for conducting group discussions were based primarily on Bohnsack's (2007, 1989) recommendations.

In addition to the collective student perspective, a third method of individual student interviews was conducted using problem-centred interviews (Witzel & Reiter, 2012; Kielblock & Lange, 2013). As a fourth method, data was collected by observing the activity of interest and taking field notes on episodes that attracted the researcher's attention.

Interview and group discussion material was audio taped and subsequently transcribed. The transcript does not reflect every single vocal expression, but it takes all of the expressed words literally. The steps of analysis are loosely based on Bohnsack's (2010, 1989) experiences and recommendations. First of all, not only the resulting text data but also the audio material of group discussions and interviews were examined to acquire an overview of the topics that were talked about. A "topic structuring" per group discussion or interview was created. The topic structuring consists basically of a chronological list of topics, each of which is specified by an appropriate topic title and a short summary of what was said about this topic. Then these documents were used to select specific passages in the transcribed material that were especially interesting on two accounts: either because relevant information was given in the passages concerning the research question or because they included very extensive and detailed descriptions. Selected passages were subsequently analysed in greater detail. The field notes were typewritten and used for an analytical illustration of the interview and group discussion analyses. This resulted in two multi-perspective case studies, each of which focused on one of the homework substitute activities.

The interpretation of the empirical material was driven by the model introduced before. The model shows that the offer (called the "homework substitute" in the following) from the teacher/pedagogue is perceived by the student body and perceived and processed by each student individually (see D>E/F, figure 2). The interpretation and presentation of the results initially reflect this process by explaining the offer (especially with regard to interviewing the member of the pedagogic staff who provided the associated activity); second, this was accomplished by revealing the students' perception in general (using group discussions with students who participated in the associated activity); and third, by illuminating the individual perception of each student (interpretation of interviews with students who participated in the asso-

ciated activity). Important additional information concerning the performance was provided by using the observation protocols. These may furnish the first clues about the transition from individual perceptions into actions (see E/F>G).

Results

Homework substitute at a Cooperative Comprehensive School

The homework substitute at the first school is called "free learning time" (German: *Freie Lernzeit*). The person providing this activity is not employed as a teacher but as supplementary pedagogic personnel. For purposes of this paper, the pseudonym "Ms. Jacobs" is used. She offered free learning time every day for about two or three school hours. The room where free learning time takes place is called the "learning workshop" (German: *Lernwerkstatt*). Folders with self-study material sorted by school subjects are available here. Students participate voluntarily in free learning time, which they attend together with other students of the same age group. The exact meaning of "voluntarily" is vague. The concept behind the learning workshop is "individual support." The students receive curriculum-based complex tasks that should be done at school (instead of at home) within the next two weeks, for example. Students who finish these tasks earlier than the classmates do extra or different tasks at the learning workshop. Ms. Jacobs pointed out that there is no regular constellation of specific students as a result. The attendance simply varies: sometimes just three students attend and sometimes there are 30.

During free learning time, students can individually choose the tasks that they work on. They can also ask which tasks they should do and complete tasks that they brought with them from their classes. Ms. Jacobs saw her duty in particular as ensuring that the students are quiet, answering questions and helping with the self-study material. She said that there are at least three fundamental ideas for the free learning time: First, students who are intellectually quicker than the rest of their class can go beyond the scope of the regular curriculum and do extra tasks. Second, students struggling with some topics in classes are encouraged during the free learning time to do extra repetitions so they can catch up with the others. And third, learning time is also good for relatively free study such as preparing for tests.

During the interview, Ms. Jacobs mentioned problems with the free learning time. She complained that learning time could not start on schedule because students did not arrive punctually and were not properly prepared. They forgot the things that they needed for the classroom work and it took time for them to finally be ready to start working.

In the first group discussion, six students confirmed that they are required to work with specific folders on the one hand and to finish tasks from classes on the other. The aspect of choosing the tasks independently and becoming actively involved in their own learning, as Ms. Jacobs also emphasised, is not clearly represented in this group discussion. However, the children have positive remarks about free learning

time in that this makes it possible for them to improve a bit in school and do better on tests. Another positive aspect is that they do not have to do any homework as a result. They also emphasised that they enjoyed attending free learning time because it ends ten minutes earlier than other classes, so they did not have to stand in a queue for lunch. (As an aside, Ms. Jacobs also explained that this is necessary because the canteen cannot deal with the rush of students if they all finish classes at the same time.)

In the second group discussion, six other students talked about how they needed the signature of the teacher to sign out of classes and in to the free learning time. The students were certain that their teachers and parents decided whether and how often they had to visit the free learning time. The group discussions suggested that free learning time was not perceived by the students as actually being "free" in terms of their individual autonomy to learn. The case seemed to be quite to the contrary.

The individual interviews with students revealed a closer look at their perception of free learning time. A few new insights were provided in comparison with the group discussions. In the face-to-face interviews, some students actually told the interviewer that it was possible to individually decide whether they wanted to do the assigned tasks or study for an upcoming test. In addition, one student explained that free learning time is only one remedial course among many others. Another student elaborated that these courses were all in different rooms and "it is totally chaotic because no one ever knows where they should go."

As shown above, the interviews revealed information about "organisational matters" such as that children are confused by the broad range of services on the one hand. If students actually struggle to comprehend the organisation of the different courses, including what to bring to them, this is not only a problem for the confused students but also creates a disturbance for those who found the right room at the right time and have the right materials with them. The field notes of the observation revealed that students do not enter the learning workshop all at once but rather in an unsteady stream of student by student.

The data suggests that students interpreted the situation not as "free" time for "free" learning but rather as a routine and clearly hierarchical learning environment. During the free learning time, students are told what to do by the pedagogue and even the teachers and parents. The field notes support this impression. From the observer's point of view, the pedagogical setting somehow mirrors regular classes: The students complete worksheets and the pedagogue walks around the room to give individual support when necessary.

However, there is at least one obstacle: Many children studied with worksheets from the folders administered by Ms. Jacobs. A student who received a single worksheet was not engaged with it for a lengthy period of time and then needed another one. Ms. Jacobs was therefore mainly busy with the folders and giving the students appropriate worksheets. As a result, she did not really have much time to accompany the students' learning progress.

Homework substitute at a German Gymnasium

Like Ms. Jacobs, the person providing the homework substitute at the second school was also not employed as a teacher. Her pseudonym is "Ms. Schrader." Ms. Schrader called her extracurricular activity "homework and learning support." Students sat together in groups and were focused on their tasks. The activity is organised across classes. There were only a few rules: "no eating, no drinking and keep it quiet."

In the first group discussion at this school, six students talked about the extracurricular time at their school. Even without being asked, they referred to the "homework support" several times. One student remembered that she attended homework support during the previous school year and that it was great because after finishing homework she was always allowed to delve into the books and have other tasks to do. Another student confirmed that this was still the good thing about homework support.

In addition, the students pointed out that Ms. Schrader is a very active learning companion. One student said:

> "Well, she doesn't sit on her chair or at her computer and check who has borrowed books. She goes around, watches for who is ready and […short passage skipped…] whether a student is ready and says 'yes I am ready.' Then she asks something like 'should I practice vocabulary with you?' And that's because of the students, that they stay quiet." (Student in group discussion)

The students emphasised that due to Ms. Schrader, a stable work climate prevailed. She always had clever ideas on how to keep the students busy with meaningful tasks. And she managed to maintain presence with all of the students even while giving individual students help with their tasks.

In the group discussion, students reflected on the differences between curricular classes and extracurricular activities. Most activities were very similar to classes. For example in the cooking activity, the pedagogue explained cooking issues and the children had to learn about them. Only homework support was seen as entirely different from the regular classes:

> Student: "Homework support is completely different because no teacher stands up in front and talks about something, saying that this works like this and that works like that. And telling us what we have to do next. Actually it is just completely different."
>
> Discussion moderator: "How is it different?"
>
> Other student: "We just work on our homework there and no teacher talks at us by saying 'you have to do it like that! No, that's wrong!'"

However, this does not mean that students want to be left alone. The case is quite to the contrary because they emphasised how important it is to have a teacher or pedagogue around who can provide help when it is needed.

One further aspect concerning homework support is that students felt that they can do homework and other tasks with each other. The presumption may have been that there would be no difference between doing homework alone and doing homework in a setting where each student must be rather quiet and focused on the individ-

ual task. Yet, being socially embedded during homework support emerged as a major issue in this group discussion.

They also discussed the fact that Ms. Schrader can get angry and be rigorous. The students in this discussion say that this was okay and even necessary to create a steady and calm work atmosphere. In the second group discussion, the other students did not reflect at this high level. They just said that Ms. Schrader is a little bit weird; initiated by this topic, the discussion was almost completely about situations in which other teachers and pedagogues behaved improperly. It was not possible to draw much additional information about homework support from this second discussion.

A closer look at homework support is provided by considering the individual interviews with students. Most commonly, the interviews reinforced the viewpoint of the first group discussion: Homework support enables students to work individually on their individual tasks. "You can do what you want" as one student stated and pointed out that he liked such flexibility and also the fact that Ms. Schrader helps students. He said that "you are not just left high and dry, especially if your parents cannot properly help you with homework or test preparation." Another student elaborated on the previously mentioned option of delving into books once the homework was done. She added that the additional reading material helps to understand specific topics. It seems that the books are not only used for fun but also for knowledge expansion.

In conclusion, the student perspective on the homework and learning support at the second school seemed to be very rich. Students emphasised working individually in homework support, attending it with the other students, receiving support and being encouraged to go beyond their regular tasks. All of this suggests that homework support significantly differs from the regular classes and other activities.

The observations and field notes support the idea, that this activity might have a significant meaning for many of the students considered here. For example, in one episode that was documented in the field notes, a student finished his arithmetic worksheet and went to Ms. Schrader to have his calculations corrected. She took the sheet and said: "Shall we check your answers with the calculator? Maybe you got everything right." The two of them fetched the calculator and Ms. Schrader started to check the student's answer. But he began a conversation with another student sitting next to Ms. Schrader's' desk. Ms. Schrader immediately gave the student his worksheet and her calculator and said: "You have to check your answers carefully." Field notes like this support the impression that Ms. Schrader very well knows how to keep students busy and quiet. If she had checked the entire worksheet by herself, the boy's conversation with another student would have disturbed the remaining students. And it seems important that she gives the students the impression that they are capable of doing even difficult tasks (e. g. correcting worksheets with the calculator) on their own.

Discussion

This study has analysed the interrelation of successful program implementation and the effectiveness of these programs. The literature search revealed that this is a common question to varying degrees. On the one hand, implementation science is at least concerned with implementation and with student outcomes as a measure of effectiveness on the other. Yet, how processes of implementation actually enhance the specific processes that lead to student outcomes is hardly considered in the literature. Therefore, the third section elaborated on the mechanisms that lead to student outcomes. This study has emphasised that the transition from the quality of the activity to the individual perception of the activity's quality appears to be crucial for further processes that lead to distinct outcomes.

Based on the empirical data of the StEG-Q project, the results demonstrate how two quite similarly implemented programs can be perceived very differently by their students. On the one hand, it is quite surprising that the individual learning support provided by the "learning workshop" in case 1 is perceived as an extension of regular classes. In outlines of the reported and observed actions of the students, they appear to engage in a kind of a sabotage of this "extra"-curricular activity (in some cases, it seems that they forgot their material on purpose, etc.). On the other hand, the individual learning support provided by the "homework and learning support" in case 2 seemed to appeal to the students because they believed that they had a real new "extracurricular" opportunity to do homework together with their classmates and friends at school.

The reported and observed actions not only support the idea that Ms. Schrader is an assertive pedagogue, but also that the students like to delve into specific tasks *together*. Although there are no measures in the StEG-Q data concerning student outcomes (see discussion of limitations), the conceptual and theoretical framework suggests that such differences in perception (and action) will result in various student outcomes (e. g. "academic achievement, learning strategies or features of the personality that are connected with learning success […] social learning, intercultural learning or a positive academic self-concept" Stecher & Maschke, 2013, p. 35).

These findings agree e. g. with the results of O'Hare (2014), who measures student perceptions of an afterschool program and the actual change in student outcomes, for example. In his quantitative analysis, he only found a "small but significant link between children's perceptions of this afterschool program and the actual change in their outcomes" (O'Hare, 2014, p. 3789). O'Hare (2014) operationalised predefined dimensions of perception: His analysis uses standardised and global measures that are not necessarily equal to what the individual personally defines or interprets as significant and relevant. This could be one explanation of why greater links between children's perceptions and outcomes are not predictable. Yet, further research is still needed to provide more conclusive evidence. One promising way to expand upon O'Hare's paper with regard to the ideas presented here might be to investigate a complex – especially an 'embedded' – mixed methods design. Having qualitative interviews systematically embedded into the quantitative analysis of the

quality-to-outcomes connection may offer a more extensive opportunity to tap the full potential of the model presented in figure 2.

In particular, case 2 showed that students interpret "doing homework at school" as a possibility of being socially embedded during "homework" time and not having to do homework at home alone. The need for relatedness seems to be a prerequisite for competence and autonomy, as it is conceptualised e. g. by self-determination theory (Deci & Ryan 1985). In this sense, it would be a desirable goal for further studies to consider self-determination theory and reflect especially on the role of relatedness. Perhaps attempts to implement *individualised* forms of learning at schools are headed in a critical direction because they lack a social component. There may be clues in this direction in the first case presented here.

However, some limitations should be mentioned at this point. These cases were specifically chosen as the most illustrative. Yet, regardless of how carefully these cases were selected and how prototypical they may represent matters of implementation, a major limitation is that this qualitative study underestimates the diversity of possible cases. A further limitation is that the two cases are utilised to illustrate the substantiation of the conceptual framework. On behalf of the detailed description of implementation issues (see chapter 2) and explanation of the model (see chapter 3), the empirical results section (see chapter 4 and chapter 5) has been kept rather brief. A paper that focusses completely on homework practices and that uses more elaborated and detailed versions of these two cases is in preparation.

There are also limitations concerning the included or available data. Some evidence in the data of StEG-Q could possibly illuminate aspects of implementation fidelity or other aspects of the implementation bridge. Including them in this paper may have made the analysis richer. This might be added in future papers when the longitudinal data of StEG-Q is completely available. The StEG-Q also has not investigated student outcomes in a narrower sense, which is why the processes leading to student outcomes cannot be fully proven in these cases. Systematically embedding the StEG-Q data into other sub-projects of the StEG (that explicitly measure different kinds of impacts and outcomes of extracurricular activities) may help to close this gap in future analyses.

This paper has created a conceptual framework that provides insight into the relatedness of program implementation and the effectiveness of these programs. In addition to the limitations, the paper has reported at least some evidence that the perceptions and actions of students should inform implementation processes. Consideration of the evidence presented here supports the conviction that a program has not been successfully implemented if it becomes a daily routine or if the implementers arrive at "new practices"; instead, is successfully implemented if it matters to the students.

References

Barbour, R. (2007). *Doing Focus Groups*. Thousand Oaks, CA: SAGE.

Blase, K. A., Van Dyke, M., Fixen, D. l., & Bailey, F. W. (2012). Implementation Science. Key Concepts, Themes, and Evidence for Practitioners in Educational Psychology. In B. Kelly & D. F. Perkins (Eds.), *Handbook of Implementation Science for Psychology in Education* (pp. 13–34). Cambridge, UK: Cambridge University Press.

Bohnsack, R. (1989). *Generation, Milieu und Geschlecht. Ergebnisse aus Gruppendiskussionen mit Jugendlichen*. Opladen: Leske + Budrich.

Bohnsack, R. (2004). Group Discussions and Focus Groups. In U. Flick, E. von Kardorff, & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 214–220). London, Thousand Oaks, CA: Sage Publications.

Bohnsack, R. (2007). *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden* (6th ed.). Opladen & Farmington Hills: Verlag Barbara Budrich.

Bohnsack, R. (2010). Documentary Method and Group Discussions. In R. Bohnsack, N. Pfaff, & W. Weller (Eds.), *Qualitative Analysis and Documentary Method in International Educational Research* (pp. 99–124). Opladen, Farmington Hills: Verlag Barbara Budrich.

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge (Mass.), London: Belknap Press of Harvard University Press.

Collier, C., & Henriksen, R. C. J. (2012). Teachers' Perceptions of a Multiple High-Risk Behavior Prevention Program and Delivery of Universal Programming. *The Qualitative Report, 17*(100), 1–19.

Davis, E. S., & Clark, M. A. (2012). Elementary School Counselors' Perceptions of Reality Play Counseling in Students' Relationship Building and Problem-Solving Skills. *Journal of School Counseling, 10*(11), 1–35.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

Deslandes, R. (2006). Designing and Implementing School, Family, and Community Collaboration Programs in Quebec, Canada. *The School Community Journal, 16*(1), 81–105.

Everhart, N. (2005). A Crosscultural Inquiry into the Levels of Implementation of Accelerated Reader and Its Effect on Motivation and Extent of Reading: Perspectives from Scotland and England. *School Library Research, 8*, 1–20.

Fiester, L. M., Simpkins, S. D., & Bouffard, S. M. (2005). Present and accounted for: Measuring attendance in out-of-school-time programs. *New Directions for Youth Development*, *105*(12), 91–107.

Fischer, N., & Klieme, E. (2013). Quality and effectiveness of German all-day schools. Results of the study on the development of all-day schools in Germany. In J. Ecarius, E. Klieme, L. Stecher, & J. Woods (Eds.), *Extended Education – an International Perspective. Proceedings of the International Conference on Extracurricular and Out-of-School Time Educational Research* (pp. 27–52). Opladen, Berlin, Toronto: Barbara Budrich Publishers.

Fischer, N., Radisch, F., Theis, D., & Züchner, I. (2012). *Qualität von Ganztagsschulen – Bedingungen, Wirkungen und Empfehlungen. Gutachten im Auftrag der SPD-Bundestagsfraktion*. Frankfurt a. M.

Greaney, M., Hardwick, C. K., Mezgebu, S., Lindsay, A. C., Roover, M. L., & Peterson, K. E. (2007). Assessing the Feasibility of a Multi-Program School-Based Intervention to Promote Physical Activity and Healthful Eating in Middle Schools Prior to Wide-Scale Implementation. *American Journal of Health Education, 38*(5), 250–257.

Grimmett, H., Rickard, N. S., & Gill, A. (2010). The Perilous Path from Proposal to Practice: A Qualitative Program Evaluation of a Regional Music Program. *Australian Journal of Music Education, 44*(2), 52–65.

Hall, G. E. (2010). Technology's Achilles Heel: Achieving High-Quality Implementation. *Journal of Research on Technology in Education, 42*(3), 231–253.

Hallenbeck, A., & Fleming, D. (2011). Don't You Want to Do Better? Implementing a Goal-Setting Intervention in an Afterschool Program. *Afterschool Matters, 13*, 38–48.

Helmke, A. (2003). *Unterrichtsqualität erfassen, bewerten, verbessern*. Seelze: Kallmeyer.

Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (3rd Ed.). Seelze-Velber: Klett/Kallmeyer.

Holtappels, H. G., Klieme, E., Radisch, F., Rauschenbach, T., & Stecher, L. (2008). Forschungsstand zum ganztägigen Lernen und Fragestellungen in StEG. In H. G. Holtappels, E. Klieme, T. Rauschenbach, & L. Stecher (Eds.), *Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der "Studie zur Entwicklung von Ganztagsschulen" (StEG)* (2nd Ed., pp. 37–50). Weinheim, München: Juventa.

Huang, J.-J., Yeh, G.-L., Tseng, C.-C., Chen, W. W., Hwu, Y.-J., & Jiang, D. D.-S. (2009). Using Organization Development Concept to Conduct Administrative Assessment of Health Promoting Schools in Taiwan – A Preliminary Study. *International Electronic Journal of Health Education, 12*(1), 89–98.

Kelly, B. (2012). Implementation Science for Psychology in Education. In B. Kelly & D. F. Perkins (Eds.), *Handbook of Implementation Science for Psychology in Education* (pp. 3–12). Cambridge, UK: Cambridge University Press.

Kielblock, S., & Lange, A. (2013). Das problemzentrierte Interview. Grundlagen und Forschungspraxis. In *Zeitschrift für Soziologie der Erziehung und Sozialisation*, *33*(4), 439–448.

Kielblock, S., & Stecher, L. (2014). Ganztagsschule und ihre Formen. In T. Coelen & L. Stecher (Eds.), *Die Ganztagsschule. Eine Einführung* (pp. 13–28). Weinheim, Basel: Beltz Juventa.

Klerfelt, A., & Haglund, B. (2014). Walk-and-Talk Conversations: a Way to Elicit Children's Perspectives and Prominent Discourses in School-Age Educare. *International Journal for Research on Extended Education, 2*(2), 119–134.

Klieme, E., & Rauschenbach, T. (2011). Entwicklung und Wirkung von Ganztagsschule. Eine Bilanz auf Basis der StEG-Studie. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher, & I. Züchner (Eds.), *Ganztagsschule:*

Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG) (pp. 342–350). Weinheim, Basel: Beltz Juventa.

Klieme, E., Lipowsky, F., Rakoczy, K., & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts "Pythagoras". In M. Prenzel & L. Allolio-Näcke (Eds.), *Untersuchungen zur Bildungsqualität von Schule* (pp. 127–146). Münster: Waxmann.

KMK (2006–2014). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland*. Retrieved from http://www.kmk.org/statistik/schule/statistische-veroeffentlichungen/allgemein-bildende-schulen-in-ganztagsform.html.

Kunter, M., Klusmann, U., Dubberke, T., Baumert, J., Blum, W., Brunner, M., et al. (2007). Linking aspects of teacher competence to their instruction. Results from the COACTIV project. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report on the DFG Priority Programme* (pp. 39–59). Münster: Waxmann.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction, 19*(6), 527–537.

McIntyre, E., Powell, R., Coots, K. B., Jones, D., Powers, S., Deeters, F., & Petrosko, J. (2005). Reading Instruction in the NCLB Era: Teachers' Implementation Fidelity of Early Reading Models. *Journal of Educational Research & Policy Studies, 5*(2), 66–102.

Miller, B. M. (2003). *Critical Hours: Afterschool Programs and Educational Success.* Quincy, MA: Nellie Mae Education Foundation.

O'Hare, L. (2014). Did Children's Perceptions of an After-School Social Learning Program Predict Change in their Behavior? *Procedia – Social and Behavioral Sciences, 116*, 3786–3792.

Ocak, Ş. (2011). Discussed Issues in Preventive Intervention Programs. *Educational Sciences: Theory and Practice, 11*(3), 1396–1402.

Olvera, N. N., Knox, B., Scherer, R., Maldonado, G., Sharma, S. V., Alastuey, L., & Bush, J. A. (2008). A Healthy Lifestyle Program for Latino Daughters and Mothers: The BOUNCE Overview and Process Evaluation. *American Journal of Health Education, 39*(5), 283–295.

Radisch, F. (2009). *Qualität und Wirkung ganztägiger Schulorganisation: Theoretische und empirische Befunde*. Weinheim und München: Juventa.

Radisch, F., Fischer, N., Stecher, L., & Klieme, E. (2008). Qualität von unterrichtsnahen Angeboten an Ganztagsschulen. In T. Coelen & H.-U. Otto (Eds.), *Grundbegriffe Ganztagsbildung* (pp. 910–917). Wiesbaden: VS Verlag.

Schwalm, J., & Tylek, K. S. (2012). Systemwide Implementation of Project-Based Learning: The Philadelphia Approach. *Afterschool Matters*, (15), 1–8.

Skaggs, G., & Bodenhorn, N. (2006). Relationships between Implementing Character Education, Student Behavior, and Student Achievement. *Journal of Advanced Academics, 18*(1), 82–114.

Stecher, L., & Maschke, S. (2013). Research on Extended Education in Germany – A General Model with All-Day Schooling and Private Tutoring as Two Examples. *International Journal for Research on Extended Education, 1*(1), 31–52.

Stecher, L., Radisch, F., Fischer, N., & Klieme, E. (2007). Bildungsqualität außerunterrichtlicher Angebote in der Ganztagsschule. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, *27*(4), 346–366.

Witzel, A., & Reiter, H. (2012). *The Problem-Centred Interview. Principles and Practice*. London: Sage.

Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C., & Misulis, K. (2011). Using the Context, Input, Process, and Product Evaluation Model (CIPP) as a Comprehensive Framework to Guide the Planning, Implementation, and Assessment of Service-Learning Programs. *Journal of Higher Education Outreach and Engagement, 15*(4), 57–84.