

Chapter 9

Effects of Student Feedback on Teaching and Classes: An Overview and Meta-Analysis of Intervention Studies



Sebastian Röhl

Abstract Based on a comprehensive literature review of student feedback intervention studies in schools, this chapter provides an overview of found effects on teachers and teaching. The first part summarizes the self-reported cognitive, affective, and motivational effects of student feedback on teachers, which can subsequently lead to behavioral changes in the classroom. In the second part, the focus is on the extent to which these behavioral changes are perceived by students. For the first time, a meta-analysis of changes in students' perceptions of teaching was carried out for the 18 existing longitudinal studies for this purpose. A small but significant positive weighted mean effect size of $d=0.21$ for students' perceived improvement of teaching quality was found, while more in-depth analyses pointed to a beneficial effect of individual support measures for teachers regarding reflection and subsequent development of teaching. Implications for further research and practical implementation of student feedback in schools are discussed.

Keywords Student feedback · Meta-analysis · Effects · Intervention studies · Teacher · Teaching development

1 Introduction

Feedback can be understood as a communicative process “in which some sender [...] conveys a message to a recipient. In the case of feedback, the message comprises information about the recipient” (Ilgen et al., 1979, p. 350). This information can be used by the recipient to improve task performance (Kluger & DeNisi, 1996) or to enable and develop learning processes (Hattie & Timperley, 2007). In the case of student feedback, the feedback recipients are teachers, who receive information on teaching from their students in class as senders. As described in the Introduction of this volume and Chap. 8 by Wisniewski and Zierer, the received feedback

S. Röhl (✉)
Institute for Educational Sciences, University of Education,
Freiburg, Germany
e-mail: sebastian.roehl@ph-freiburg.de

should contain useful and meaningful information for the given teacher. As a first step, the feedback could therefore have positive cognitive and possibly also affective and motivational effects on the teacher. Subsequently, this could lead to changes in teacher behavior, thus promoting development and improvement of teaching and professionalism. This in turn could lead to a more positive perception of teaching by students.

This overview chapter follows this process and is based on a comprehensive literature review of studies dealing with student feedback as an intervention for the improvement of the teaching quality of fully trained teachers. In the first part, findings on teacher-reported effects from student feedback are summarized. The second part contains a meta-analysis of findings of longitudinal student feedback intervention studies, which almost exclusively examined changes in teaching and classes from the perspective of students in secondary schools. Remarkably, no studies could be found which were conducted in grades one to four.

This chapter complements Chap. 11 by Göbel et al., which describes the use of student feedback in the context of the first and second phases of teacher training. In Chap. 12, Schmidt and Gawrilow describe how student feedback can be used to improve the cooperation between teachers and students. Furthermore, teachers' productive use of student feedback depends on various individual and situational characteristics, and this is described by Röhl and Gärtner in Chap. 10 and in the Introduction of this volume.

2 Self-Reported Effects of Student Feedback on Teachers

Whether a feedback message leads to visible changes in the recipient's behavior depends on the effects of the feedback message on the recipient—in this case the teacher. Therefore, this part offers an overview of literature on self-reported effects of student feedback on teachers. For the teacher obtaining feedback, student feedback can have effects at different levels (see *processes and effects of student feedback model (PESF)* in the Introduction of this book). Here, a distinction can be made between affective, cognitive, and behavioral effects, which in turn are related to motivational processes.

Regarding *cognitive* effects of obtaining student feedback, several studies reported an increasing amount of reflection by teachers on their actual practice due to aspects of teaching quality included in the used feedback questionnaires (Gärtner & Vogt, 2013; Göbel & Neuber, 2019; Mandouit, 2018). As a result of the feedback received, teachers express an improvement regarding their understanding of how students perceive their teaching and classes (Gage, 1963; Thorp et al., 1994; Wyss et al., 2019). Furthermore, student feedback can help teachers to find students' misconceptions about learning (Mandouit, 2018). Subsequently, teachers identified possible areas for improvement (Barker, 2018; Gaertner, 2014). As a side effect, the first-time use of student feedback can lead to a more positive attitude towards this instrument

(Brown, 2004; Campanale, 1997; Gaertner, 2014), although opposite effects such as a higher skepticism have also been observed (Dretzke et al., 2015).

On the *affective* level, many teachers experience emotions of happiness and curiosity during the feedback reception and reflection, especially if the feedback is perceived as positive (Villa, 2017). Other teachers reported emotions of anger due to feedback perceived as negative, or sadness due to helplessness regarding a possible improvement of their own teaching (Brown, 2004; Gärtner & Vogt, 2013; Villa, 2017).

Both cognitive and affective effects can impact motivational processes and lead to changes on the *behavioral* level. Teachers expressed that they paid more attention to identified improvement areas during preparation and teaching, sometimes resulting in a self-perceived improvement (Balch, 2012; Gaertner, 2014; Rösch, 2017). In addition, some teachers planned to participate in relevant professional training programs (Balch, 2012). Another behavioral outcome is the discussion about feedback received and teaching with the corresponding class, which was seen by many teachers as an important further source of information about their own teaching and a common ground for changing teaching practices (Gaertner, 2014; Thorp et al., 1994). In addition, teachers mentioned changes in their behavior before obtaining student feedback. While reflecting on the feedback questionnaire, they prepared the lessons in which the instrument was to be used more carefully, in line with the questionnaire's quality criteria (Balch, 2012; Rösch, 2017).

3 A Meta-Analysis of Longitudinal Studies on the Teaching-Related Effects of Student Feedback Interventions

Without a doubt, it is desirable that positive effects of student feedback are not only reported by teachers, but that they also become evident in student perceptions and learning achievement. Based on the process model of Student Feedback on Teaching (SFT, see the Introduction to this book), this process can only be achieved if several conditions are met. First of all, the students have to report back that there is a need for improvement. This must be perceived and accepted by the teacher in the feedback reports. Furthermore, it is necessary that the teacher creates a desire for change or sets goals and then pursues them. Subsequently, a teacher's behavioral change should improve students' learning processes—and students have to perceive this behavioral change—before a positive effect of student feedback on teaching and classes becomes visible.

While intervention studies on the use of students' achievement data for the instructional development in schools also focus on student achievement (e.g. Keuning et al., 2019; van der Scheer & Visscher, 2018) or the improvement of teachers' instructional skills (van der Scheer et al., 2017), the overwhelming focus of investigations into student feedback has been on the effects of student perception of teaching behavior. In

the literature review performed here, one single study (Novak, 1972) was found which additionally analyzed several audio-recorded lessons before and after the student feedback intervention on changes in teacher behavior. The findings of this study pointed to significantly lower proportions of teacher talk and lectures during the lessons following repeated reception of student feedback. Regarding possible effects of student feedback interventions on students' motivation, findings of a single study (Tozoglu, 2006) indicated a small positive effect ($d = 0.289$), but only for teachers who received enhanced support for interpreting feedback and teaching development. No effect was found for teachers who received only student feedback mean scores without any support. In a dissertation study, Kime (2017) measured students' achievement scores in the context of teaching evaluations based on student ratings, comparing a group of teachers receiving student feedback only with another group which carried out additional peer coaching on the feedback received. Contrary to Kime's expectations, analysis could not prove a significant effect on achievement scores for the peer coaching condition. However, a comparison with teachers who did not receive student feedback was not possible due to the lack of an appropriate control group. With regard to the question of the extent to which primary school pupils perceive an improvement in the quality of teaching, a study by van der Scheer (2016) resulted in no changes in pupils' rating of teaching quality during a data-based decision-making intervention, whereas pupils' learning achievement significantly improved. Research concerning effects on students' learning achievement, comparing teachers receiving student feedback with non-receivers, is still absent.

While in the field of university and college teaching some meta-analyses of effects of students' mid-term feedback on classes already exist (e.g. Cohen, 1980; L'Hommedieu et al., 1990; Penny & Coe, 2004), a meta-analysis regarding effects in schools is still pending. The meta-synthesis regarding feedback by Hattie (2009), which resulted in $d = 0.73$, and also a recent and thorough meta-analysis of the underlying primary studies with a lower effect size of $d = 0.48$ (Wisniewski et al., 2020), mainly include feedback from teachers to students, with the exception of three meta-analyses of effects of student feedback in higher education. For the context of higher education, Cohen's (1980) meta-analysis of 17 intervention studies resulted in an effect size of $d = 0.20$ on students' end-of-semester ratings of classes for providing mid-term feedback to university teachers. If the feedback is accompanied by further measures such as individual consultation, this effect increases to an average of $d = 0.64$. Penny and Coe (2004) found an average effect size of $d = 0.69$ for student feedback augmented with peer and expert consultation in their analysis of 11 intervention studies. The analysis of 28 studies by L'Hommedieu et al., (1990) resulted in $\Delta = 0.34$. Uttl et al. (2017) conducted a meta-analysis of 51 studies on the relation between student evaluation of teaching ratings and student learning achievement. The results indicated no significant overall correlation. In order to close this research gap in the field of primary and secondary schools, a meta-analysis is now presented here, which includes student feedback intervention studies while surveying changes in students' perception of teaching quality.

3.1 Measures and Methods

3.1.1 Literature Search

For this overview, a comprehensive literature search using the terms “student feedback”, “pupil feedback”, and “self-evaluation” was conducted in the databases ERIC, PsycInfo, Scopus, Web of Science, ProQuest, and OpenDissertations. As most of the studies found focus on student feedback in higher education, the search was limited to publications which did not contain this keyword. In a second step, articles with a theoretical or practical focus were excluded. Next, only intervention studies which reported pre- and post-measures were selected. In addition, some non-catalogued studies mentioned in scientific articles on student feedback were found. More details about the studies included can be seen in Part III.

3.1.2 Study Coding

Regarding possible moderators, most of the different study characteristics are explicitly reported, such as the existence of a control group, the number of feedback reports, the duration of the treatment, and the publication type. For the level of provided support for the participating teachers (see below), a coding was conducted by two trained raters. The inter-rater agreement was high ($\rho = 0.85$, $p < 0.001$), and in the subsequent discussion a consensus was reached on the different opinions.

3.1.3 Effect Size Calculation and Analysis

The dependent variable in this meta-analysis is the student-perceived change in the quality of teaching. As the included studies use different questionnaires for student feedback, single scales or constructs are not comparable across the studies. Therefore, in order to achieve comparability of the effects, it was decided to calculate the arithmetic mean of all reported effect sizes included in each study for the students’ perception of teaching as an overall effect.

Effect sizes are calculated using Cohen’s d with groups-size-adjusted standard deviation (σ_{pooled} , Morris & DeShon, 2002). Effect size variances were estimated following Lipsey and Wilson (2001, pp. 44–49). If available, d was estimated using the reported means and standard deviations of pre- and post-measurements on teacher level. Otherwise, available t , F , and χ^2 statistics were used.

In this meta-analysis, longitudinal studies with and without a control group design are included. This led to some problems in the estimation of comparable effect sizes and variances:

- (a) Several studies without control groups didn’t include the standard deviations of the measurements and the correlation between the pre- and post-test scores. While comparable effect sizes can be estimated without this information using

reported t - or F -values (Lipsey & Wilson, 2001), the variances of the effect sizes can only be estimated if the standard deviations or correlations are available. For this meta-analysis, several solutions were considered. The most conservative approach would be to assume no correlation between the two measurement time points, which would lead to a strong overestimation of variances. However, many studies report quite high consistency of student ratings on teaching quality over time (e.g. Polikoff, 2015; Rowley et al., 2019). In addition, the calculation of the correlation between teachers' pre- and post-measures using available data from two studies (Bartel, 1970; Ditton & Arnold, 2004) results in values of $r > 0.73$. Therefore, following the suggestions of Borenstein et al. (2009), we assumed a lower limit of $r = 0.70$ for the estimation of effect sizes variances.

- (b) Many studies with a control group design showed a moderate decrease of control groups' student ratings on teaching quality between the measurement time points (Buurman et al., 2018; Gage, 1963; Nelson et al., 2015; Tacke & Hofer, 1979; Tuckman & Oliver, 1968). For the studies using a control group design, this tendency is already considered in the estimation of effect sizes. However, assuming that this effect is also evident in the treatment group, this could lead to an underestimation of the strength of the effect in designs without the control group. Therefore, possible moderator effects regarding the design of the study are included in our analyses.

Because of the heterogeneity of treatment and design characteristics of the included studies, random-effect models appeared to be more suitable than fixed-effect models for this meta-analysis (Borenstein et al., 2009). For the estimation of assumed moderator effects of study and treatment characteristics, separate mean weighted effect sizes and confidence intervals for every subgroup were estimated (Borenstein et al., 2009). Regarding continuous study characteristics such as the number of feedback reports and the intervention duration, the studies were split at the median. Estimation of the overall and moderator effect sizes and confidence intervals was done using the package *metafor* (Viechtbauer, 2010) in *R* (R Core Team, 2019). In addition, as three studies included several effect sizes by different intervention groups, a sensitivity analysis was conducted with regard to bias due to possible dependencies (Hedges et al., 2010). This revealed that the resulting biases are about $d = 0.0001$, and therefore negligible.

Analysis on possible outliers or influential studies was conducted. We chose to use Cook's distance (Cook & Weisberg, 1982) test statistics for residual heterogeneity when each study is removed in turn (Viechtbauer, 2010), and the distribution of weights of the included studies as indicators.

3.2 Characteristics of Included Studies

In the literature review, 18 longitudinal studies with student feedback treatments published between 1960 and 2019 were identified (see Table 1). The design of these studies is experimental or quasi-experimental. Thus, all studies include at least one pre- and one post-measurement of students' perception of teaching quality, but not all of them provide a control group comparison. Seven of the studies were conducted in the USA, three more took place each in Australia and Germany, two in the Netherlands, and one each in Great Britain, Turkey, and Austria.

All studies utilized questionnaires which were mainly based on closed questions or rating scales. The research teams carried out the counting and provided a feedback report to the teachers. One study used a digital smartphone-based feedback system for this purpose (Bijlsma et al., 2019). All included studies were conducted in grade 5–13. While five interventions were limited to exactly one grade level, the other studies involved teachers from different levels. Three interventions also continued to restrict the subject matter for a better comparability of the classes. Novak (1972) focused on biology teachers, Rösch (2017) on physics, and Bijlsma et al. (2019) on mathematics.

The findings on the effects of a student feedback intervention on changes in teaching behavior perceived by students are heterogeneous in the studies. While two studies show clearly negative treatment effects (Bennett, 1978, $d = -0.30$; Knox, 1973, $d = -0.24$),¹ most studies report effects ranging from $d = 0.1$ to $d = 0.5$.

Furthermore, some studies instruct teachers to focus on only one to three areas for improvement in subsequent classroom development (Fraser & Fisher, 1986; Fraser et al., 1982; Nelson et al., 2015; Thorp et al., 1994). However, information on which aspects were selected by teachers for improvement is only available for the three case studies. As expected, results show the highest improvements in the targeted areas (up to $d = 0.8$), whereas the other scales do not change. Another study (Mayr, 1993, 2008) examined only individual areas of teaching which had been agreed with the teachers. However, as there is a complete lack of such information for all other studies, the individual prioritization of certain areas by individual teachers cannot be considered in this meta-analysis, and so we used the average effect sizes of all scales in each study. This also means that the average overall effects of all included scales are smaller than the reported bigger improvements in some selected scales.

The *sample size* differs greatly between the studies. Whereas some have reported case studies with single teachers (Fraser & Fisher, 1986; Fraser et al., 1982; Thorp et al., 1994) or one team of five teachers (Mandouit, 2018), the other studies used sample sizes ranging from $N = 10$ to $N = 508$ teachers. Also, the duration of the intervention varied between the studies from one month to one year, with an average of $M = 3.06$ months. During these periods, a different *number of feedbacks* were reported to the teachers. In most of the studies, the last feedback report was used

¹Noteworthy, both studies were conducted by persons from the school administration. To what extent the negative effect can be explained by possible refusal attitudes of subordinate teachers cannot be clarified here due to the small number of studies.

Table 1 Studies included in the meta-analysis

No.	Study	Country, types of school and grades	N teachers		Duration in months	Number of feedbacks	Level of support	Effect size <i>d</i>	Publication type
			Treatment	Control					
<i>Intervention studies with control group</i>									
1	Gage et al. (1960, 1963)	USA, secondary, 6	86	90	1.5	1	Low	0.17	Journal
2	Tuckman and Oliver (1968)	USA, high and vocational, 10–13	186	100	3	1	Low	0.30	Journal
3	Bartel (1970) • Feedback Group • Counselling Group	USA, secondary, 7–12	14	14	3	1	Low	0.08	Dissertation
			14	14	3	1	High	0.47	
4	Bennett (1978)	USA, high, 9–12	19	18	2	3	Low	−0.30	Dissertation
5	Tacke and Hofer (1979)	Germany, secondary, 7–10	22	22	1	1	Low	0.24	Journal
6	Tozoglu (2006) • Feedback only Group • Enhanced Feedback Group ^a	Turkey, high, 9–12	20	20	5	1	Low	0.61	Dissertation
			20	20	5	1	High	1.61	
7	Nelson et al. (2015)	USA, middle, 6–8	16	15	1	1	Medium	0.17	Journal
8	Buurman et al. (2018)	Netherlands, vocational, 11–13	116	126	12	1	Low	0.11	Report
9	Bijlsma et al. (2019)	Netherlands, secondary, 9	28	25	4	4–17	Low	0.13	Journal

(continued)

Table 1 (continued)

No.	Study	Country, types of school and grades	N teachers		Duration in months	Number of feedbacks	Level of support	Effect size <i>d</i>	Publication type
			Treatment	Control					
<i>Intervention studies without control group</i>									
No	Study	Country, types of school and grades	N Teachers		Duration in months	Number of feedbacks	Level of support	Effect size <i>d</i>	Publication type
10	Novak (1972)	USA, high, 10	18		1	5	Low	0.19	Dissertation
11	Knox (1973)	USA, high, 10-12	20		3	2	Medium	-0.24	Dissertation
12	Fraser (1980, 1982)	Australia, elementary & secondary, 6&7	2		1.5	1	High	0.21	Journal
13	Mayr (1993, 2008)	Austria, lower secondary track	27		2	1	High	0.83	Report
14	Thorp et al. (1994)	Great Britain, comprehensive, 8	1		1	1	High	0.65	Journal
15	Ditton and Arnold (2004)	Germany, secondary and high, 5-13	46		8	1	Low	-0.09	Journal
16	Bell and Aldridge (2014) • Reflection Group ^a • Action Research Group	Australia, secondary, "upper grades"	508 52		2 2	1 1	Low High	0.17 0.42	Journal
17	Rösch (2017)	Germany, middle and grammar, 9	10		3	4	Low	0.12	Dissertation
18	Mandouit (2018)	Australia, secondary, n/a	5		3	3	High	0.47	Journal

^aExcluded after outlier analysis

as post-measure of changes in the student perceived teaching quality or teacher behavior. Therefore, for comparability reasons, we counted the number of student feedback reports before the last measurement. Whereas 11 studies reported only one student feedback measurement to the teachers, the other studies obtained and reported feedback up to five times. A special case in point is the study by Bijlsma et al. (2019), where teachers could use the smartphone app to obtain feedback as often as they wanted. The frequency varied between 4 and 17 feedback measurements, with an average of 6.7 for these teachers.

The studies reported here differ also in the manner and amount of *support* provided for the feedback interpretation and subsequent developmental processes. In line with the meta-analysis results from higher education described above (Cohen, 1980; Penny & Coe, 2004), findings on teachers' use of students' achievement data pointed out that solely providing data rarely leads to subsequent changes in teaching (Schildkamp et al., 2015). Thus, it seems to be important to consider this characteristic of the interventions. Furthermore, three of the included studies analyzed different treatment conditions (Bartel, 1970; Bell & Aldridge, 2014; Tozoglu, 2006). One part of the teachers received written feedback without further instructions, while the other part received additional reflection impulses and counseling. All three studies showed significantly more positive effects for the latter condition. For this reason, the effects of these different treatments are reported as two separate effect sizes for each of these studies in the meta-analysis. During the coding process of the support by the raters it became apparent that the following three levels of support can be distinguished:

- Low level of support: General training of student feedback use. This support level includes introductory explanations and training on the use of student feedback before the start of the intervention. These were partly given in written form but also in face-to-face sessions. Also, studies which do not contain explicit descriptions of this topic were assigned to this level. If the information is missing, we assume that the participating teachers were appropriately instructed in the use of the feedback questionnaires and reports.
- Medium level of support: Individual reflection support for the feedback received. This more intense kind of support includes an individualized feedback report with the special marking of possible developmental areas. This occurs in written form and also in face-to-face meetings.
- High level of support: Individual support for subsequent teaching development. Furthermore, some interventions also included ongoing advice on the subsequent development processes through individual or group consultations, counseling, or professional learning communities.

A further distinguishing feature of the studies is the *type of publication*. While the findings of some studies were published in peer-reviewed journals, others were only available as reports or university theses and required a high search effort to find them. If only studies from scientific journals are included in meta-analyses, this easily leads to a so-called "publication bias", since these usually contain higher effects and more significant findings than those not included in such journals (Lipsey & Wilson,

2001). An analysis on differences of effects between publication types could provide indications on whether a publication bias also exists for this research field (Borenstein et al., 2009). Of course, this leaves the question unanswered to what extent further studies exist which could not or cannot be found.

3.3 Results of the Meta-Analysis

A first estimation of the mean weighted effect size using all 21 effect sizes found in a random-effects model resulted in $d = 0.23$ ($p < 0.001$, 95%-C.I.: 0.13–0.33). Analyses of influential studies pointed to an overweight of the reflection group in the study of Bell and Aldridge (2014) because of the exceptional sample size. In addition, analysis of the residual heterogeneity led to the exclusion of the enhanced feedback group from Tozoglu (2006) due to outlier characteristics of this subsample.

For the remaining 19 effect sizes, the estimation of the overall mean weighted effect size led to $d = 0.21$ ($p < 0.001$) with a 95% confidence interval of $0.11 < d < 0.32$. The effect sizes with confidence intervals of all included studies are plotted in Fig. 1.

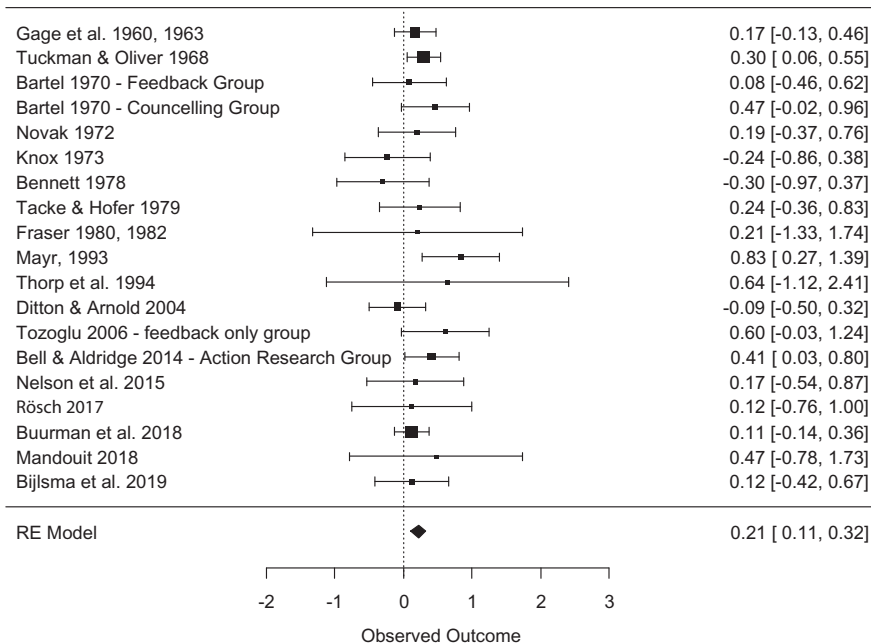


Fig. 1 Forest plot of effect sizes and 95% confidence intervals of included studies and the mean weighted effect size

The inspection of the heterogeneity test statistics ($Q(18) = 16.62, p = 0.549$) reveals that the homogeneity of the effect size is statistically sufficient (Lipsey & Wilson, 2001).

3.3.1 Moderator Analysis

The resulting mean effect sizes and 95% confidence intervals of the subgroups split along the moderator variables are presented in Table 2. In line with the relatively small numbers of studies found, confidence intervals overlap mostly between the different subgroups.

The only study characteristic which turned out to be a significant moderator is the level of support. Treatments with a high level of individual support for reflecting on feedback and teaching development (level 3) showed a significantly higher effect size ($d = 0.52, p = 0.010$) than studies with a medium or low supportive level. Contrary to the assumptions, no significant differences were found between the effect sizes of studies including control groups and studies without ($d = 0.21$ vs. $d = 0.24$). The differences (presumed as considerable) between studies with only one or with more feedback reports ($d = 0.25$ vs. $d = 0.01$) were not statistically relevant ($p = 0.123$). The same applies to the differences regarding the treatment duration of the

Table 2 Analysis of moderator effects regarding study and treatment characteristics

	<i>n</i>	<i>d</i>	95%-CI	<i>p</i> ^a
Design of study				
with control group	10	0.21	[0.01; 0.81]	0.817
without control group	9	0.24	[-0.04; 0.52]	
Level of support ^b				
low	11	0.16	[0.04; 0.28]	0.089
medium	2	-0.06	[-0.53; 0.40]	0.235
high	6	0.52	[0.26; 0.77]	0.010
Number of feedback reports				
1	13	0.25	[0.13; 0.36]	0.123
2 or more	6	0.01	[-0.26; 0.29]	
Duration of treatment				
1–2 months	9	0.27	[0.08; 0.46]	0.461
3–12 months	10	0.18	[0.05; 0.32]	
Publication type				
Peer reviewed journal	10	0.22	[0.08; 0.37]	0.824
Thesis or report	9	0.22	[0.00; 0.44]	

^aSignificance of moderator, ^bDummy-coded

intervention and to whether the studies are published in scientific journals or only accessible as theses or reports.

4 Conclusion and Discussion

In this overview chapter, findings of a comprehensive literature review on effects of student feedback interventions in schools were presented. In the first step, effects on teachers were summarized from the literature found. Regarding cognitive effects, studies reported reflective thinking processes on teachers' own perceptions and goals of teaching—initiated by feedback reports and also by questionnaire topics—which could lead to an identification of areas for improvement. In addition, a fostering effect on teachers' understanding of students' perception of teaching and learning processes was observed. Both positive (happiness, joy) and negative (sadness, feelings of helplessness) affective reactions are found with regard to the feedback received. Cognitive and affective processes can result in motivational effects, which could change teachers' behavior in classes. According to teachers' self-reports, these behavioral changes are apparent in a more intense preparation of lessons and a stronger perception and control of one's own actions in class, if they consider the feedback points as critical. Furthermore, teachers initiated discussions with students about the received feedback and the improvement of teaching and collaboration within the school class.

In a second step, this chapter examined whether and to what extent behavioral changes by teachers were perceived by the students. To answer this question, the first meta-analysis of effects of student feedback interventions on student-perceived teaching quality in schools was conducted, including 18 studies with 19 effect sizes. Using a random-effects model, a weighted mean effect size of $d = 0.21$ was found. Although this effect seems to be relatively small, it is significant and lies in a similar range to meta-analyses from student feedback use in higher education (Cohen, 1980; L'Hommedieu et al., 1990). Furthermore, it should be noted that these analyses were based on all the teaching characteristics assessed by the students, but teachers often focused only on specific areas for improvement. For the target areas, the case studies in particular showed considerably greater effects. In addition, the effect sizes varied to a considerable extent between the different scales of teaching dimensions used in the larger studies.

Additional moderator analysis showed an increase in the effect size to $d = 0.52$ for additional individual support, which is also in line with findings for college and university teachers (Penny & Coe, 2004). Other moderator analyses showed no significant effects. This emphasizes the important impact of providing appropriate teacher support for the feedback-related teaching development process, whereas other structural treatment characteristics play no or only a minor role. However, there were indications that further studies should pay particular attention to the number of feedback reports provided in longer-term studies.

Considering the findings of the first part of this chapter on the teacher-reported effects of feedback, the teacher's perception processes and reactions are the “needle's

eye” for improving teaching. Therefore, support for teachers using student feedback should aim at facilitating a constructive cognitive processing of feedback and accompanying affective reactions, so that teachers can develop action alternatives and thus the motivation for change is fostered.

As a limiting factor for the meta-analysis presented, it should be noted that only relatively few studies were found. This reduces the power of the analyses of possible moderators. However, the similarity of the findings presented here to meta-analyses from higher education points toward validity of these results, together with the fact that there is no indication of a publication bias or design effect of the included studies. This chapter thus provides evidence for the effectiveness of student feedback as a tool for improving the quality of teaching perceived by students. It provides a comprehensive overview of the effects on teachers which have so far only been considered in isolation in studies. Furthermore, an extensive literature review and meta-analysis of intervention studies on student feedback in schools was presented for the first time.

Simultaneously, there are various implications for further research on the effects of student feedback in schools:

- With one exception, only intervention studies which measure changes in teaching based on student perceptions or teacher self-reports have been conducted to date. Hence, there is an urgent need for studies which measure changes in teaching using other methods such as video analysis or student achievement.
- The findings of this study point to the importance of additional support to teachers for productive use of student feedback. However, it has not yet been controlled to what extent the supporting measures would have the same positive effect if, for example, self-assessments of teachers were used instead of student feedback.
- Studies should include which areas of improvement have been identified by teachers and analyze these effects separately.
- In addition, there is also a lack of studies which focus both on teachers’ reflection processes on feedback together with the subsequent changes in teaching, perceived by students or external observers.

For the practical use of student feedback for teaching development in schools, this meta-analysis also results in several implications. Most importantly, the findings emphasize the need for support for teachers on using student feedback. This does not only concern the subsequent lesson development, but also support for the interpretation of feedback reports, dealing with accompanying emotions, identification of improvement areas, and how to work on them. This can for example take place through coaching and supervision, but also in collegial settings such as professional learning communities.

Additionally, when planning the implementation of student feedback in schools, there is a need to consider organizational characteristics which are beneficial for constructively dealing with feedback, as presented in Chap. 10 by Röhl and Gärtner in this volume.

Acknowledgements The author wants to thank Prof. Dr. Martin Schwichow and Prof. Dr. Wolfram Rollett from the University of Education Freiburg for their valuable suggestions and support for this study.

References

*Studies included in the meta-analysis are marked with an asterisk**

- Balch, R. T. (2012). *The validation of a student survey on teacher practice*. Dissertation. Vanderbilt University, Nashville, Tennessee.
- Barker, S. (2018). *Student voice to improve instruction: Leading transformation of a school system*. Digital Commons @ ACU.
- *Bartel, B. W. (1970). *The effectiveness of student feedback in changing teacher classroom image*. University of Minnesota.
- *Bell, L. M., & Aldridge, J. M. (2014). Investigating the use of student perception data for teacher reflection and classroom improvement. *Learning Environments Research*, 17, 371–388. <https://doi.org/10.1007/s10984-014-9164-z>.
- *Bennett, C. R. (1978). *A developed and field-tested experiment to test the effect of increased student feedback on specific teacher performance behaviors*. Dissertation. Iowa State University, Ames, IA.
- *Bijlsma, H. J. E., Visscher, A. J., Dobbelaer, M. J., & Veldkamp, B. P. (2019). Does smartphone-assisted student feedback affect teachers' teaching quality? *Technology, Pedagogy and Education*, 28, 217–236. <https://doi.org/10.1080/1475939X.2019.1572534>.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.
- Brown, S. L. (2004). *High school students' perceptions of teacher effectiveness: Student ratings and teacher reflections*. Dissertation. University of New Mexico, Albuquerque, NM.
- *Buurman, M., Delfgaauw, J. J., Dur, R. A. J., & Zoutenbier, R. (2018). *The effects of student feedback to teachers: Evidence from a field experiment* (Tinbergen Institute Discussion Paper). Amsterdam & Rotterdam.
- Campanale, F. (1997). Autoévaluation et transformations de pratiques pédagogiques. *Mesure Et Évaluation En Éducation*, 20(1), 1–24.
- Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education*, 13, 321–341. <https://doi.org/10.1007/BF00976252>.
- Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. Chapman and Hall.
- *Ditton, H., & Arnold, B. (2004). Wirksamkeit von Schülerfeedback zum Fachunterricht [Effectiveness of student feedback on teaching]. In J. Doll & M. Prenzel (Eds.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (pp. 152–172). Waxmann.
- Dretzke, B. J., Sheldon, T. D., & Lim, A. (2015). What do K-12 teachers think about including student surveys in their performance ratings? *Mid-Western Educational Researcher*, 27(3), 185–206.
- *Fraser, B. J., & Fisher, D. L. (1986). Using short forms of classroom climate instruments to assess and improve classroom psychosocial environment. *Journal of Research in Science Teaching*, 23, 387–413. <https://doi.org/10.1002/tea.3660230503>.
- *Fraser, B. J., Seddon, T., & Eagleson, J. (1982). Use of student perceptions in facilitating improvement in classroom environment. *Australian Journal of Teacher Education*. <https://doi.org/10.14221/ajte.1982v7n1.3>.
- Gaertner, H. (2014). Effects of student feedback as a method of self-evaluating the quality of teaching. *Studies in Educational Evaluation*, 42, 91–99. <https://doi.org/10.1016/j.stueduc.2014.04.003>.

- *Gage, N. L. (1963). A method for “improving” teacher behavior. *Journal of Teacher Education*, 14, 261–266. <https://doi.org/10.1177/002248716301400306>.
- Gärtner, H., & Vogt, A. (2013). Wie Lehrkräfte Ergebnisse eines Schülerfeedbacks verarbeiten und nutzen [How teachers process and use results of student feedback]. *Unterrichtswissenschaften*, 41(3), 252–267.
- Göbel, K., & Neuber, K. (2019). Lernende geben Rückmeldungen zum Unterricht: Potenziale der Nutzung von Schülerfeedback und deren Bedingungen [Learners give feedback on teaching: potentials of using student feedback and their conditions]. *Friedrich Jahresheft*, 37, 48–49.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement* (1st ed., Educational research). Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112. <https://doi.org/10.3102/003465430298487>.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1, 39–65. <https://doi.org/10.1002/jrsm.5>.
- Ilgen, D. R., Fisher, C. D., & Taylor, S. M. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), 349–371.
- Keuning, T., Geel, M., Visscher, A., & Fox, J.-P. (2019). Assessing and validating effects of a data-based decision-making intervention on student growth for mathematics and spelling. *Journal of Educational Measurement*, 56, 757–792. <https://doi.org/10.1111/jedm.12236>.
- Kime, S. J. M. (2017). *Student evaluation of teaching: can it raise attainment in secondary schools? A cluster randomised controlled trial*. Dissertation. Durham University, Durham. <http://etheses.dur.ac.uk/12267/>.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
- *Knox, J. T. (1973). *Student feedback as a means of improving teacher effectiveness*. Dissertation. Wayne State University, Detroit, MI.
- L’Hommedieu, R., Menges, R. J., & Brinko, K. T. (1990). Methodological explanations for the modest effects of feedback from student ratings. *Journal of Educational Psychology*, 82, 232–241. <https://doi.org/10.1037/0022-0663.82.2.232>.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Applied Social Research Methods Series, Vol. 49). SAGE.
- *Mandouit, L. (2018). Using student feedback to improve teaching. *Educational Action Research*, 13, 1–15. <https://doi.org/10.1080/09650792.2018.1426470>.
- *Mayr, J. (1993). *Mitarbeit und Störung im Unterricht: Beschreibung und Evaluierung eines Konzepts zur Verbesserung pädagogischen Handelns* [Collaboration and classroom disturbance: description and evaluation of a concept to improve pedagogical action]. Linz.
- *Mayr, J. (2008). Forschungen zum Führungshandeln von Lehrkräften: Wie qualitative und quantitative Zugänge einander ergänzen können [Research on teacher leadership: How qualitative and quantitative approaches can complement each other]. In F. Hofmann, C. Schreiner, & J. Thonhauser (Eds.), *Qualitative und quantitative Aspekte: Zu ihrer Komplementarität in der erziehungswissenschaftlichen Forschung* (pp. 321–341). Waxmann.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105–125. <https://doi.org/10.1037/1082-989X.7.1.105>.
- *Nelson, P. M., Ysseldyke, J. E., & Christ, T. J. (2015). Student perceptions of the classroom environment: Actionable feedback to guide core instruction. *Assessment for Effective Intervention*, 41, 16–27. <https://doi.org/10.1177/1534508415581366>.
- *Novak, J. H. (1972). *A study of the effects of the use of a pupil response instrument on the behaviors of biological science teachers*. Final Report. Pittsburgh, PA.

- Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74, 215–253. <https://doi.org/10.3102/00346543074002215>.
- Polikoff, M. S. (2015). The stability of observational and student survey measures of teaching effectiveness. *American Journal of Education*, 121(2), 183–212.
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- *Rösch, S. (2017). *Wirkung und Wirkmechanismen von regelmäßigem Schülerfeedback in der Sekundarstufe: Eine explorative Untersuchung im Physikunterricht* [Effect and impact principles of regular student feedback in secondary education: An exploratory study in physics classrooms]. Dissertation. Universität Basel, Basel.
- Rowley, J. F. S., Phillips, S. F., & Ferguson, R. F. (2019). The stability of student ratings of teacher instructional practice: Examining the one-year stability of the 7Cs composite. *School Effectiveness and School Improvement*, 30, 549–562. <https://doi.org/10.1080/09243453.2019.1620293>.
- Schildkamp, K., Poortman, C. L., & Handelzalts, A. (2015). Data teams for school improvement. *School Effectiveness and School Improvement*, 27, 228–254. <https://doi.org/10.1080/09243453.2015.1056192>.
- *Tacke, G., & Hofer, M. (1979). Behavioral changes in teachers as a function of student feedback: A case for the achievement motivation theory? *Journal of School Psychology*, 17, 172–180. [https://doi.org/10.1016/0022-4405\(79\)90025-6](https://doi.org/10.1016/0022-4405(79)90025-6).
- *Thorp, H. S., Burden, R. L., & Fraser, B. J. (1994). Assessing and improving classroom environment. *School Science Review*, 75, 107–113.
- *Tozoglu, D. (2006). *Effects of student ratings feedback on instructional practices, Teaching Effectiveness, and Student Motivation*. Florida State University.
- *Tuckman, B. W., & Oliver, W. F. (1968). Effectiveness of feedback to teachers as a function of source. *Journal of Educational Psychology*, 59, 297–301. <https://doi.org/10.1037/h0026022>.
- Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22–42. <https://doi.org/10.1016/j.stueduc.2016.08.007>.
- van der Scheer, E. A. (2016). *Data-based decision making put to the test*. Dissertation. University of Twente, Twente.
- van der Scheer, E. A., Glas, C. A. W., & Visscher, A. J. (2017). Changes in teachers' instructional skills during an intensive data-based decision making intervention. *Teaching and Teacher Education*, 65, 171–182. <https://doi.org/10.1016/j.tate.2017.02.018>.
- van der Scheer, E. A., & Visscher, A. J. (2018). Effects of a data-based decision-making intervention for teachers on students' mathematical achievement. *Journal of Teacher Education*, 69(3), 307–320.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.
- Villa, L. A. L. (2017). *Teachers taking action with student perception survey data*. Arizona State University.
- Wyss, C., Raaflaub, M., & Hüsler, N. (2019). Schülerrückmeldungen zur Förderung der Partizipation in der Schule [Student feedback to promote participation in school]. In S. Hauser & N. Nell-Tuor (Eds.), *Mündlichkeit* (1st ed., pp. 181–209, Sprache und Partizipation im Schulfeld, Vol. 6). hep.
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, 3087. <https://doi.org/10.3389/fpsyg.2019.03087>.

Sebastian Röhl has been an Academic Assistant in the Department of Educational Science at the University of Education Freiburg and is currently a Postdoctoral Researcher in the Institute of Education at Tübingen University (Germany). Before that he worked for more than 10 years as a grammar school teacher, school development consultant, and in teacher training. Among other

areas, he conducts research in the fields of teaching development and teacher professionalization through feedback, social networks in inclusive school classes, as well as teachers' religiosity and its impact on professionalism. In addition, he is the director of an inservice professional master's study program for teaching and school development.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

