Teacher Collaboration in German Schools: Do Gender and School Type Influence the Frequency of Collaboration Among Teachers?

Julio G. Mora-Ruano1*, Markus Gebhardt2 and Eveline Wittmann1

1 TUM School of Education, Technical University of Munich, Munich, Germany, 2 Faculty of Rehabilitation Sciences, Technical University of Dortmund, Dortmund, Germany

According to the literature, collaboration among teachers can be regarded as a subject that can positively influence numerous aspects of the school. However, factors influencing the extent to which collaboration takes place have received less attention. For example, gender differences are usually not examined. We carry out a secondary analysis of the representative PISA 2012 German teacher sample to measure the state of teacher collaboration through three different collaboration forms namely instruction-, project-, and organization related. Because of the stratification of the German school system in secondary education, variations by school type should be taken into account. We conducted a MANCOVA to investigate the differences and effects that gender and school type of the lower secondary education have on the frequency with which teachers collaborate. Results show that women collaborate minimally more than men, that the higher the formal education of a school, the less frequently collaboration occurs and that the overall level of collaboration is weak. There is no interaction effect between these two variables. Suggestions for future research as well as implications are presented.

Keywords: teacher collaboration, Germany, PISA 2012, MANCOVA, gender, school type, age, experience

INTRODUCTION

Collaboration among teachers is consistently regarded as an influential component of education that can have a positive impact on several variables such as students’ outcome (Goddard et al., 2007) and job satisfaction (Johnson et al., 2012). However, although the importance of collaboration has been heavily emphasized throughout the years, with interest steadily growing; research on the subject is notably scarce, especially about gender differences, despite the fact that the profession of teaching differs significantly between women and men in everyday life (Koch-Priewe, 1997).

Studies investigating gender differences and its influence at the teacher level in Germany are unfortunately limited as they tend to focus more on school differences and no insight in gender differences is provided (cf. Helmke and Jäger, 2002; Bos et al., 2004; Kanders and Rösner, 2006; Steinert et al., 2006). Considering that the German education system comprises a well-defined hierarchy of school types, and each school type offers something specific to both students and teaching staff, it is also relevant to examine mean differences at the school type.
We use representative data from PISA 2012 (Prenzel et al., 2015) to examine the frequency in which teachers collaborate. Apart from some brief mentions in the official PISA reports, we could not find any studies that have used PISA data to investigate the current state of collaboration among teachers in German schools, despite the fact that in Germany since PISA 2003, data through the “National Lehrerfragebogen” (questionnaire for teachers) has been gathered.

**REVIEW OF THE LITERATURE**

Numerous studies of school effectiveness and school development have conceptualized teacher collaboration as a central construct for school quality, school effectiveness and student achievement (cf. Steinert et al., 2006; Goddard et al., 2007; Williams, 2010; Bondorf, 2013). Researchers and professional organizations portray collaboration as an element lying at the very foundation of effective professional development for teachers. It has been pointed out as an extremely important factor for the development of the school and its members, showing a strong positive association with teachers’ self-efficacy and job satisfaction (Ahlgrimm, 2010; OECD, 2016) and is recognized in Germany as a crucial factor for the development of the school (Avenarius et al., 2003). Collaboration has also been placed at the process level of models of school quality and school efficiency (Sheerens, 1990; Ditton, 2000). Hattie (2015) claims that “the greatest influence on student progression in learning is having highly expert, inspired and passionate teachers and school leaders working together to maximize the effect of their teaching on all students in their care” (p. 2). In the context of integrated education of persons with disabilities, collaboration has also been considered a fundamental criterion for the quality of the inclusive education itself (Gebhardt et al., 2013). Morse (2000) suggested that collaboration is an educational reform imperative: “educators will recognize they are not alone in searching for new modes of human exchange. The fact is, this quest for a new way of human exchange is endemic in the social order… rejecting collaboration is not an option” (p. xi).

**Gender**

Schümer (1992), in a study “use of media in class” conducted in Baden-Württemberg, Berlin, Hessen and Nordrhein-Westfalen, found that women collaborate more than men and suggests that specific gender forms to exercise the teaching profession are the reason behind this finding. Ulich (1996) came to a similar conclusion and writes that due to their socialization, women seemingly have better interaction skills than men, who in turn behave more competitively at school. In another study (Maag Merki et al., 2010) showed that for women several aspects of the collaboration among teachers (e.g., enjoying common work, exchange of expertise, interest in collaboration with colleagues, recognition of school management and authority) are more important than for men. Soltau (2011) found significant mean differences among four forms of collaboration (exchange, student related exchange, synchronization and co-construction) between women and men but no correlation with other variables, such as self-concept. Richter and Pant (2016) also found that women collaborate more than men in two of the three forms of collaboration used in their analyses. However no explanation or interpretation for this result is given.

**School Type**

Research in Germany shows frequently that teachers in Gymnasium (grammar schools) collaborate less than their peers in other schools types of the German school system. A comparison of German and English teachers within the DESI Study 2006 indicates that teachers at Gesamtschulen (comprehensive schools) collaborate the most in the preparation for lessons while Gymnasium teachers collaborate the least (DESI Konsortium, 2006). In a study conducted in all school types in Bremen, Soltau (2007) found that among the three forms of teacher collaboration proposed by Gräsäl et al. (2006), primary school teachers have the highest level of collaboration, whereas Gymnasium teachers have the lowest.

In another study conducted in Hessen and the Swiss Kanton Zurich, Steinert et al. (2006) proposed five different levels of teacher collaboration (presented here in ascending form, the first being the lowest level): fragmentation, differentiation, coordination, interaction and integration. Their results were similar to the ones already presented; teachers in the Gymnasium find themselves mostly in the first two levels while non-Gymnasium teachers were at the other end of the spectrum. The mean differences between both groups were small but significant. The authors suggest that due to a higher subject differentiation and eligibility criteria of the students, teachers in Gymnasium are less challenged to integrate collaboration in their activities. Finally, the DESI Study 2006 indicates that there is a greater frequency of collaboration among German teachers in terms of joint lesson preparation than joint lesson implementation. Across the literature, the overall conclusion is that the level of collaboration in German schools is low.

**General Remarks and Difficulties**

It should not go unnoticed that, despite its importance for processes within the school and although the empirical evidence supports its effectiveness, collaboration is a remarkably complex process that requires time, trust, social skills and conflict management, just to name a few inherent challenges. These factors do not appear spontaneously and need to be developed and nurtured. As Henning (2013) writes: "Teachers working together in the best interest of all students can be a powerful thing, but also difficult" (p. 121).

Although in essence, teaching is a notably interpersonal activity, its own unusual features (e.g., the uncertainty provided by the social complexities in a classroom) make it self-centered (Evans, 2012). One can say that it has a tendency to isolate teachers; however, these feelings of isolation can be reduced through collaboration (Reeves et al., 2017). Professional discussions among teachers are often considered lacking or non-existent, due to the tradition of working alone. Previous research suggests that teachers do not look for opportunities to share with colleagues; they avoid communicating in ways that make others feel imposed upon (Ertesvåg, 2011). This
suggests that there is a sort of self-imposed isolation. Altrichter and Eder (2004) point out that it protects teachers from negative criticism and the transgression of autonomy, which is one of the most treasured elements of teachers. In other words, it seems that autonomy has a conflicting relationship with collaboration. However, Vangrieken et al. (2017) argue that “autonomy in itself does not have a negative relationship with collaboration” (pp. 305). According to the Kaiser Institute (2015) “autonomy without collaboration ultimately ends in utter isolation […] collaboration without autonomy ultimately ends […] in diffusion.” Teachers, fundamentally do not want to influence others’ work or have their own work influenced (McLaughlin and Talbert, 2001; Hargreaves, 2005; Levine and Marcus, 2008), resulting in a limitation of the possibilities of mutual learning and reflection processes (Altrichter and Eder, 2004).

Yet, the reasons that make the development and implementation of collaboration so difficult lie not only in the individual but also in the administration. As Fimian (1982) writes: “a lack of administrative support reinforces teachers’ perceptions that they are totally on their own. When individuals start to feel their classroom has become their entire world, they actually can become divorced from the real world” (p. 102). To put it simply, the difficulties lie within both the institutions and the teachers themselves to accomplish successful collaboration practices.

Furthermore, it must be stated that collaboration does not always result in innovative teacher learning or generate new knowledge or practices (Meirink et al., 2010). Additionally, it does not necessarily contribute to successful student learning in “collaborative groups” in classroom settings (Barron, 2003), and therefore, the effectiveness of teacher collaboration builds on a number of factors such as leadership (Graham, 2007).

All these findings suggest that many changes, from the individual to the organizational level, are required in order to develop, implement and build up successful collaborative practices among teachers, and even under the best of circumstances, these changes are difficult to achieve and may take several years to accomplish (McLeskey and Waldron, 2006; Fullan, 2007). Results from the study of Fovargue (2008) showed, for example, a tendency of teachers to not fully apply the strategies or knowledge acquired during their professional development activities (as cited in Morgan, 2010).

Therefore, because of the inherent benefits and challenges collaboration among teaching staff is comprised of, it is of paramount importance to devote a great deal of attention to the state of collaboration in our schools, so the development and implementation of collaborative practices can be a reality that works for schools, teachers and students.

Definitions

Across the literature, there seems to be a consensus about the lack of a uniform and generally accepted definition of collaboration and more specifically the lack of a precise definition suited to the specific needs at the teacher collaboration level. Aldorf (2016) sustains that due to the absence of a precise definition the respective research fields examine the same concept from different points of view and conceptualize it heterogeneously. One possible reason for this could be that researchers attempt to keep definitions as simple as possible to avoid the use of too specific or complicated terms. On the one hand, this constitutes an advantage, for the concept will be easily grasped; on the other hand, this can make the definitions too vague for a specific given field. Kelchtermans (2006) writes:

“Teacher collaboration as a term and an object of educational research seems obvious and self-evident in its meaning. Yet, even a quick look at the literature shows that the term is far from being unequivocal. Further definition and specification are necessary in order to properly discuss the issue” (p. 220).

The lack of a uniform, consistent definition often imply that the conceptualization, the study and the findings of certain forms of teacher collaboration are confounded with other forms (Soltau, 2007). This approach leads to an incomplete description of the term, especially when a specific context is given. This leaves aside the huge range of prerequisites and objectives that a successful collaborative practice among educators requires. Consequently, at the end of this section we will introduce our own definition of teacher collaboration. However, it is necessary to clearly define collaboration as a general construct first. For example, collaboration and cooperation are commonly interchanged; however, they mean something inherently different. In a nutshell, cooperation can be defined as “working together to accomplish shared goals” (Smith, 1995), whereas collaboration is “a method that implies working in a group of two or more to achieve a common goal, while respecting each individual’s contribution to the whole” (McInerney and Robert, 2004, p. 205). Mattessich and Monsey (1992) draw a clearer difference between the two terms:

Cooperation is characterized by informal relationships that exist without any commonly defined mission, structure or planning effort. Information is shared as needed, and authority is retained by each organization so there is virtually no risk. Resources are separate as are rewards. Collaboration connotes a more durable and pervasive relationship. Collaborations bring previously separated organizations into a new structure with full commitment to a common mission […] risk is much greater because each member of the collaboration contributes its own resources and reputation. Resources are pooled or jointly secured, and the products are shared (p. 42).

Across the literature, definitions from several perspectives can be found, and they are adapted to somehow fit into the collaboration among teachers. These perspectives vary from general dictionary-type definitions to more specific, like psychoanalytic (Deutsch, 1949), occupational and organizational psychology (Piepenburg, 1991; Spieß, 2004) and political education (Reinhardt, 2000), to name a few. A more specific definition for the school context is the pedagogic-oriented one by Esslinger (2002):

“Collaboration is understood as an activity between two or more persons which is initiated and implemented with the aim of increasing the effectiveness of the work and the satisfaction at work. Collaboration requires a common target perspective. The prerequisites for this are an existing potential for structural and functional coordination and work fields” (p. 62).
Another specific definition directed at schooling is the one by Bauer and Kopka (1996): “by collaboration, we mean the goal-oriented collaboration of at least two teachers, who try to work on common tasks more effectively, efficiently and with more human satisfaction than anybody alone could do” (p. 143).

We integrate and expand these approaches, defining teacher collaboration as a voluntary activity between two or more teachers who, based on relational trust and respect, through collaborative leadership and school administration, coordinate efforts, reconcile different approaches and exchange ideas and materials in order to increase teaching effectiveness as well as affective and cognitive job satisfaction. We consider this a suitable definition of teacher collaboration as it takes into account the school context, its core function of teaching, as well as a cultural and a micro-political perspective; aspects that Kelchtermans (2006) consider to be the foundation of teacher collaboration. It comprises personal, organizational and school specific characteristics that a successful collaboration practice between teachers requires. Concerning our study, this definition provides a clear and appropriate framework for the three forms of collaboration the IPN built from the Bosker and Hendriks (1997) questionnaire. Outside the scope of this study, it can be used to gain a deeper understanding of previously conceptualized forms of collaboration, like the ones proposed by Gräsel et al. (2006) and additionally, through this definition, a further differentiation of the concept of collaboration can be made. Future studies could provide a clearer framework taking into account all the necessary aspects of collaboration among teachers.

Research Questions and Hypotheses

The aim of our study is to investigate, by analysing the PISA 2012 data from the teacher questionnaire in Germany, whether school type and gender have an effect on collaboration and to what extent the frequency of various forms of collaboration among the different school types as well as between genders varies. We use age and experience in years as covariates. We want to answer the following research questions:

1. To what extent does gender influence the frequency of teacher collaboration?
2. To what extent does the school type influence the frequency of teacher collaboration?
3. Are there significant interaction effects between school type and gender on the frequency of teacher collaboration?

We have shown that only a handful of studies have investigated gender differences. In the light of this apparent lack of research it is sensible to investigate mean differences among women and men, including the main effects of gender on various forms of collaboration. We expect that, like shown in previous studies, women collaborate more than men and that gender has significant main effects on the frequency of preparation. Given the previous findings and the additional fact that the emphasis for the preparation of Gymnasium teachers lies more on the academic content of their areas rather than on pedagogical theory, we expect that Gymnasium teachers collaborate significantly less than their peers in other types of schools. Finally, we do not expect to find a significant interaction effect between school type and gender.

METHODS

Design

In this study, we carry out a secondary analysis by using the representative PISA 2012 data from the German National Questionnaire for Teachers (questionnaire by Bosker and Hendriks, 1997). Through the IPN (Leibniz Institute for the Education of Natural Sciences and Mathematics), three different types of teacher collaboration were formed from the items namely:

Instruction-related (IRC) which refers to preparatory aspects of teaching like the exchange of teaching materials or the preparation of individual lessons; it promotes a common culture of teaching and the common development of didactic and methodological skills. Project-related (PRC) which covers joint implementations of curricula in the classroom from the planning of entire lessons to the preparation of written exams. Finally, organization, performance and problems related (ORC) which focuses on efforts to efficiently use the individual students’ learning time by means of e.g., discussion how to help students depending on their performance.

A multi-sample, also known as cluster design, was used in order to ensure the representativeness of the sample. In this sampling design, schools are first selected, and then within each selected school, classes, students or teachers are randomly selected. For the specific school situation in Germany, some precise characteristics have to be taken into account, such as the structure in the different federal states. With help from statistical authorities, a complete list of all schools in Germany was created, in which potential students at the age of 15 could be found. The age “15 years” is defined by the period of birth from January 1, 1996 to December 31, 1996 for the purposes of PISA 2012. This list includes the basic population from which the schools were randomly drawn. The total population was divided into the 16 federal states and the number of schools by means of the current number of 15-year-old students as well as of the existing general educational school types. This division is called stratification, and each state is a stratum (layer). In addition to the 16 federal states, the basic population consists of two more strata: the vocational schools as well as schools for children with special needs. This ensures that all federal states and all schools are covered in a proportion that meets the requirements of representativeness (Sälzer and Prenzel, 2013).

Participants

Of the 2,084 schoolteachers across Germany who took part in the study, 1040 were female (49.9%), 917 were male (44%), and 127 were not specified/missing (6.1%), and they ranged in age from 25 to 71 years (M = 46 years, SD = 11.4). They come from all school types: Hauptschule (main schools), 231 (11.1%); MBG (schools with several education programmes), 305 (14.6); Realschule (secondary schools), 397 (19.0%); Gesamtschulen (comprehensive schools); 251 (12.0%); Gymnasium (grammar schools), 753 (36.1%); Berufliche Schulen (vocational schools),
47 (2.3%) Sonder- and Förderschule (schools for children with special needs), 98 (4.7%); and 2 (0.1%) not specified/missing.

A critical remark should be noticed: although in PISA 2012, data from Berufsschule and Sondern- und Förderschule has been gathered, the proportion of these schools, attended by 15-year-olds in Germany, is too small to make a representative statement about them. Samples from these schools were collected only to secure the overall picture of the achievements of the 15-year-olds at schools in Germany and not to carry out analyses per school type; PISA makes statements about the total population of 15-year-olds attending a school in various educational systems worldwide (Sälzer and Prenzel, 2013). For this reason and given that we want to investigate differences at the school type, samples from these two school types were excluded from our analysis.

In the subsequent sections, the German denominations for the school types will be used.

Setting
The questionnaire took approximately 30-min to complete. Participants were required to use black ink. Additionally, teachers were asked to make their best estimation when an exact answer to a question could not be given (for example, some questions are specifically related to mathematics teaching at grade 9, but participants normally teach a second subject). Participants were informed that the collection and handling of the data was confidential. Responses of all teachers were grouped to form overall and average scores, making it impossible to identify individual schools or individual teachers. Through the questionnaire, detailed data is requested on demographic information, questions about school and education, basic orientation, collaboration among teaching staff, perception of the management’s work, continuing education and training, questions about mathematical teaching and work with parents.

Materials
In PISA 2012, the three forms of collaboration (IRC, PRC and ORC) are investigated through question number 21 (see Appendix). The items comprised in this question have been used since PISA 2006. Information on each question is provided in Appendix. To this date, PISA 2012 has been the only year in which detailed information about collaboration was gathered through these questions. PISA 2009, for example, used only one question to measure the levels of collaboration among teachers, and it was suitable only for German language teachers, making its use quite limited. Prior to writing this article, no data for PISA 2015 have been published for use outside official reports.

Analysis
All analyses were conducted using the software package SPSS 24. Missing data were due to missing responses on item level. The missing rate was <3% and therefore considered unproblematic (Graham et al., 2003); nevertheless, given that multivariate methods require complete data, a multiple imputation analysis was performed in order to account for these missing values. The multiple imputation (MI) technique introduced by Rubin (1987) was selected over other techniques to handle missing data, such as listwise or pairwise deletion, because it is considered by many statisticians to be the “gold standard” for handling missing data (Treiman, 2009, p. 185). Multiple imputation is a procedure that deals properly with statistical analyses given the uncertainty caused by the presence of missing data (Manly and Wells, 2015). The automatic method was selected because the software selects the best method available depending on the pattern of missing values, namely monotone or fully conditional specification.

A two-way MANCOVA (gender * school type with age and experience as covariates) was run to investigate the differences between the three forms of collaboration comprised in question 21 among teachers in German schools and to identify main and interaction effects. These effects were measured through Partial η². Cohen (1988) and Vacha-Haase and Thompson (2004) defined effect sizes as small (≥0.01), medium (≥0.06), and large (≥0.14) in behavioral or social science studies. In order to make these comparisons, we condensed the items belonging to each dimension into a “super-scale” using their means. Age and experience was used as control variable.

In our analyses, we carry out a comparison between Gymnasium and all other school types as a single condensed case (AOS), because Gymnasium is the only type of school that remains consistent across all German Federal States (Bundesländer). All types of schools included in the analysis belong to Lower Secondary Education ( Sekundarstufe I).

Originally in the questionnaire, question 21 used a scale consisting of six Likert categories (1 = never, 2 = once a year, 3 = several times per year, 4 = every month, 5 = every week, 6 = every day). However, given that some of the categories overlapped with each other (e.g., every week, every day) the scale was condensed into four categories that are more meaningful with the following values: 1 = never, 2 = once a year, 3 = several times per year/month, 4 = every week/every day. A summary of the general question can be seen in Appendix.

RESULTS

Testing of Assumptions
Prior to conducting the MANCOVA, a series of tests were performed in order to test the assumptions required to run such an analysis. A series of Pearson correlations were conducted between all of the dependent variables in order to test the MANCOVA assumption that the dependent variables would be correlated with each other in the moderate range (Meyers et al., 2006). As shown in Table 1, a meaningful pattern of correlations was observed amongst all the dependent variables, suggesting the appropriateness of a MANCOVA.

<p>| TABLE 1 | Correlation between collaboration forms. |</p>
<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Instruction-Related Collaboration</td>
<td>–</td>
<td>0.688*</td>
</tr>
<tr>
<td>2. Project-Related Collaboration</td>
<td>0.629*</td>
<td>0.626*</td>
</tr>
<tr>
<td>3. Organization-Related Collaboration</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*All correlations were significant at the p < 0.01 level.
Normal distribution of the dependent variables, as recommended by Ghasemi and Zahediasl (2012), was assessed visually and through the Shapiro-Wilk-Test. The results of the Kolmogorov-Smirnov-Test were disregarded owing its low power. Although two of the three scales across the school types were significant at \( p < 0.05 \), and therefore considered not to be normally distributed, this is a test extremely sensitive to minor deviations of normality, especially in large sample sizes. According to the Central Limit Theorem, as the sample size becomes larger, the distribution of means calculated from repeated sampling will approach normality. Furthermore, an inspection of the Q-Q Plots revealed that the data were in fact normally distributed. Considering all this, we carried on with the analyses. Reliability of the scales were measured through Cronbach’s \( \alpha \) and all of them resulted in an acceptable value: IRC \( \alpha = 0.68 \); PRC \( \alpha = 0.79 \) und ORC \( \alpha = 0.74 \). Additionally, a Box’s M value of 30.70 was associated with a \( p \)-value of 0.032, which was interpreted as non-significant based on the guideline and therefore, showed that the covariance matrices between the groups are equal.

Equality of variances were assessed through the Levene’s Test and although its \( F \)-test suggested that the variances associated with the IRC and PRC scales were not homogenous, an examination of the standard deviations revealed that none of the largest value was more than four times the size of the corresponding smallest, suggesting that the MANCOVA would be robust in this case (Howell, 2009). Furthermore, Field (2013) points out that “when the sample size is large, small differences in group variances can produce a Levene’s test that is significant” (p. 150).

**Multivariate Effects**

Significant multivariate effects were found for all of the IVs (Table 2). There was a significant but minimal gender effect on the combined dependent variables, \( F_{(3,1663)} = 3.435, p < 0.05 \), Wilks’ \( \Lambda = 0.994 \), partial \( \eta^2 = 0.006 \). Additionally, there was a small and statistically significant school type effect on the combined dependent variables, \( F_{(3,1663)} = 30.705, p < 0.05 \), Wilks’ \( \Lambda = 0.948 \), partial \( \eta^2 = 0.052 \). Age, which was used as covariate yielded a non-significant effect, \( F_{(3,1663)} = 2.204, p > 0.05 \), Wilks’ \( \Lambda = 0.996 \), partial \( \eta^2 = 0.004 \). Experience, which was the other covariate, yielded a non-significant effect, \( F_{(3,1663)} = 0.879, p > 0.05 \), Wilks’ \( \Lambda = 0.995 \), partial \( \eta^2 = 0.002 \). No significant multivariate interaction effects for school type and gender were detected \( F_{(3,1663)} = 0.861, p = 0.46 \), Wilks’ \( \Lambda = 0.998 \), partial \( \eta^2 = 0.002 \).

Follow-up univariate ANOVAs were run to examine the effect of the dependent variables on the collaboration scales.

**Research Question 1: Gender**

With the exception of the ORC scale, which resulted in a significant but minimal effect (Table 3), the IRC and PRC scales yielded non-significant results. ORC, \( F_{(1,1665)} = 9.984, p < 0.05 \), partial \( \eta^2 = 0.006 \). IRC, \( F_{(1,1665)} = 2.129, p > 0.05 \), partial \( \eta^2 = 0.001 \) and PRC, \( F_{(1,1665)} = 2.278, p > 0.05 \), partial \( \eta^2 = 0.002 \). All simple pairwise comparisons were run with a Bonferroni adjustment applied and showed that in all forms of collaboration, women collaborate more than men (Figure 1) being the difference bigger at the ORC scale which yielded a mean difference of 0.092.

The marginal means for the IRC score were 2.798 (SE = 0.020) for women and 2.755 (SE = 0.020) for men. The marginal means for the PRC score were 2.255 (SE = 0.021) for women and 2.203 (SE = 0.021) for men. Pairwise comparisons yielded non-significant results on these two scales. The marginal means for the ORC score were 2.742 (SE = 0.020) for women and 2.651 (SE = 0.020) for men. This was the only scale where pairwise comparisons yielded significant results. There was a statistically significant mean difference between the women and men (0.092, 95% CI [0.35, 0.149], \( p < 0.001 \)).

**Research Question 2: School Type**

All scales revealed significant main effects (Table 4). There was a significant but minimal main effect for the IRC scale: \( F_{(1,1665)} = 7.9, p < 0.05 \), partial \( \eta^2 = 0.005 \). Both the PRC and the ORC scales yielded a small and statistically significant main effect: PRC, \( F_{(1,1665)} = 68.37, p < 0.05 \), partial \( \eta^2 = 0.039 \); ORC, \( F_{(1,1665)} = 33.63, p < 0.05 \), partial \( \eta^2 = 0.020 \). All simple pairwise comparisons were run with a Bonferroni adjustment applied (Figure 2) and showed that in all forms of collaboration, Gymnasium teachers collaborate significantly less than their peers in other school types (Figure 2) being the difference bigger at the ORC scale which yielded a mean difference of 0.092.

The marginal means for the IRC score were 2.816 (SE = 0.018) for the AOS and 2.737 (SE = 0.022) for the Gymnasium. The mean difference was significant (0.081, 95% CI [0.024, 0.135], \( p < 0.01 \)). The marginal means for the PRC score were 2.255 (SE = 0.019) for the AOS and 2.203 (SE = 0.023) for the Gymnasium. The mean difference was significant (0.247, 95% CI [0.189, 0.306], \( p < 0.01 \)). The marginal means for the ORC score were 2.780 (SE = 0.018) for the AOS and 2.613 (SE = 0.022) for the Gymnasium. The mean difference was significant (0.167, 95% CI [0.110, 0.223], \( p < 0.001 \)).

**Research Question 3: Interaction Effects**

There was no significant interaction effect between school type and gender on any of the dependent variables: IRC, \( F_{(1,1665)} = 2.3, p > 0.05 \), partial \( \eta^2 = 0.001 \); PRC, \( F_{(1,1665)} = 2.0, p > 0.05 \), partial \( \eta^2 = 0.001 \); ORC, \( F_{(1,1665)} = 0.84, p > 0.05 \), partial \( \eta^2 = 0.001 \).

---

**Table 2** | Multivariate effects.

<table>
<thead>
<tr>
<th>Effect</th>
<th>( \Lambda )</th>
<th>( F )</th>
<th>( p )</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.994**</td>
<td>3.435</td>
<td>0.016</td>
<td>0.006</td>
</tr>
<tr>
<td>School type</td>
<td>0.948**</td>
<td>30.71</td>
<td>0.000</td>
<td>0.052</td>
</tr>
<tr>
<td>Age</td>
<td>0.966*</td>
<td>2.20</td>
<td>0.086</td>
<td>0.004</td>
</tr>
<tr>
<td>Experience</td>
<td>0.998*</td>
<td>0.879</td>
<td>0.451</td>
<td>0.002</td>
</tr>
<tr>
<td>School * gender</td>
<td>0.998*</td>
<td>0.861</td>
<td>0.461</td>
<td>0.002</td>
</tr>
</tbody>
</table>

*non-significant; **significant at \( p < 0.05 \).
TABLE 3 | Significant univariate effects for gender (significant at $p < 0.001$).

<table>
<thead>
<tr>
<th>DV</th>
<th>$\eta^2$</th>
<th>$df$</th>
<th>$df$ error</th>
<th>$F$</th>
<th>Gender</th>
<th>Means</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORC</td>
<td>0.006</td>
<td>1</td>
<td>1665</td>
<td>9.984</td>
<td>Female</td>
<td>2.742</td>
<td>2.702–2.783</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Male</td>
<td>2.651</td>
<td>2.611–2.691</td>
</tr>
</tbody>
</table>

The covariates were calculated by an age value of 46.1 and experience (in years) with a value of 18.1.

DISCUSSION

Through this article, we wanted to explore the state of teacher collaboration in Germany, as measured by the national questionnaire for teachers in PISA 2012. Specifically, we wanted to know how strong or weak the frequency of teacher collaboration is, taking into account school type and gender. That is, we investigate to what extent school type and gender influence the overall collaboration of the three postulated collaboration forms at school with regard to teaching: instruction related (IRC), project related (PRC) and organizational, problems and performance related (ORC).

Previous studies suggest that the magnitude of the collaboration between teachers in German schools is relatively small (Gräsel, 2008), that the teachers in Gymnasium collaborate less than their peers in other school types (Helmke and Jäger, 2002; Kullmann, 2008), that women collaborate more than men (Schümer, 1992; Soltau, 2011) and their influence may be dependent upon the type of collaboration.

Gender showed a significant but minimal influence on both the combined dependent variables and its separated forms. Differences between female and male teachers have shown that, in all three forms of collaboration, women collaborate more often than men; however, these differences yielded mixed results in terms of its significance, but across this factor, the effects on the collaboration frequency were rather small or non-existent at all. Consequently, we take great care to interpret the results from this variable.

In our analysis, only the school type has a significant effect on both the overall collaboration and all its separated forms. Additionally, the instruction-related collaboration (IRC) had marginally the highest mean value over the other
two forms of collaboration. The project-related collaboration (PRC) was the type of collaboration used least. Moreover confirming previous research, through the representative PISA sample, teachers in Gymnasium have the smallest frequency of collaboration on all the scales. Overall, the differences were small. Age, which was used as a covariate in our analysis, gave mixed results regarding its significance, its effect on the IRC and PRC scales being significant but minimal. Experience in years, which was also used a covariate, resulted in non-significant results for all the three forms of collaboration.

Following these findings, it is our opinion that more effort to increase teacher collaboration should be made at the school level. A successful collaborative practice in a specific type of school might not work in another. Hence, local adaptation is required in order to increase collaborative practices among teachers and to be able to gain a deeper understanding of its processes and effects.
LIMITATIONS AND FUTURE RESEARCH

Some limitations of the study warrant attention. First, this study used data from PISA 2012; therefore, a newer dataset might shed more light into the state of teacher collaboration in Germany. Second, given that the focus of this article is to investigate the state of collaboration in German schools as measured by PISA 2012, we conducted no exploratory nor confirmatory analyses, in order to prove the factorial validity of the instrument; consequently we used the three types of collaboration constructed by the IPN. Future studies which want to perform analysis with causal conditions, might examine its factorial validity. Third, this instrument has only been applied to a German sample of teachers; thus, future studies need to confirm that the instrument can be used in other cultural. Finally, due to legal impediments, a comparison between Bundesländer was not possible to conduct, although empirical studies conducted previously in selected Bundesländer, suggest a certain amount of variation notwithstanding that teachers in Gymnasium generally collaborate the least.

In this article, no causal relationships were tested (e.g., student outcomes, job satisfaction); future research could test for causal relationships in order to see the effects that teacher collaboration has on other variables such as students’ outcomes. Overall, more evidence is needed to draw conclusions about the importance, the state and the perception of teacher collaboration.

CONCLUSION

Research has consistently placed teacher collaboration as an important factor for the development of schools, the job satisfaction of teaching staff and the improvement of student outcomes. However, its implementation can be extremely challenging, as coordination and involvement of the entire school community (principals, teachers, students, etc.) is required.

Two major conclusions can be made from this study. One is that the overall level of collaboration in German schools is not strong, no matter the type of collaboration: instruction, project or organizational. The second conclusion is that Gymnasium teachers collaborate less than their colleagues in other school tracks on all these types of collaboration, meanwhile, results remain inconclusive with regard to gender differences in collaboration, which are minimal at best.

Collaboration is a very complex form of interaction, and therefore, straightforward and plain definitions may unintentionally suggest that the concept itself is simple. However, collaboration is far away from being an easy endeavor. For this reason, we have presented what we consider a more accurate definition that suits the specific requirements for teacher collaboration. At its heart, collaboration encompasses several activities and many variables that educators, not only as professionals but as community members, must develop and enforce in order to be successful collaborators. In depth research, covering the different aspects of the concept of collaboration put forward in this article may lead to more nuanced results.

ETHICS STATEMENT

Permission to access and use the data for scientific purposes was granted through the German Research Data Center (FDZ) at the Institute for Educational Quality Improvement (IQB). As per OECD guidelines and German national regulations (KMK) no new ethics approval was required. The authors did not have access to identifiable information.

AUTHOR CONTRIBUTIONS

JM-R drafted the manuscript, wrote the literature background, performed and interpreted the statistical analyses. MG and EW contributed to the framing of the article, edited the manuscript and discussed the data analyses.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program.

REFERENCES


**APPENDIX**

**Question 21**

<table>
<thead>
<tr>
<th>Construct: Collaboration (forms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item text: How often do you use the following forms of collaboration with colleagues?</td>
</tr>
<tr>
<td>Data source: National questionnaire for teachers (biology, chemistry, physics, natural science or mathematics).</td>
</tr>
<tr>
<td>Answer format: Rating scale</td>
</tr>
<tr>
<td>Literature/Remarks: Bosker and Hendriks (1997). Adaptation by the IPN.</td>
</tr>
<tr>
<td>Number of items: 17</td>
</tr>
<tr>
<td>Categories:</td>
</tr>
<tr>
<td>1 = never</td>
</tr>
<tr>
<td>2 = once in a year</td>
</tr>
<tr>
<td>3 = several times in a year</td>
</tr>
<tr>
<td>4 = every month</td>
</tr>
<tr>
<td>5 = every week</td>
</tr>
<tr>
<td>6 = every day</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item ID.</th>
<th>Item text</th>
</tr>
</thead>
<tbody>
<tr>
<td>21i01</td>
<td>exchange of teaching materials$^a$</td>
</tr>
<tr>
<td>21i02</td>
<td>exchange of examination questions$^a$</td>
</tr>
<tr>
<td>21i03</td>
<td>preparation of individual lessons$^a$</td>
</tr>
<tr>
<td>21i04</td>
<td>joint planning of entire lessons or projects$^b$</td>
</tr>
<tr>
<td>21i05</td>
<td>planning interdisciplinary lessons$^b$</td>
</tr>
<tr>
<td>21i06</td>
<td>joint implementation of lessons$^b$</td>
</tr>
<tr>
<td>21i07</td>
<td>testing new teaching ideas and methods$^b$</td>
</tr>
<tr>
<td>21i08</td>
<td>peer observation$^b$</td>
</tr>
<tr>
<td>21i09</td>
<td>reconciliation of dealing with homework$^c$</td>
</tr>
<tr>
<td>21i10</td>
<td>interdisciplinary discussion of student performance$^c$</td>
</tr>
<tr>
<td>21i11</td>
<td>preparation of replacement hours$^c$</td>
</tr>
<tr>
<td>21i12</td>
<td>follow-up lessons$^a$</td>
</tr>
<tr>
<td>21i13</td>
<td>monitoring and advising new teachers$^a$</td>
</tr>
<tr>
<td>21i14</td>
<td>joint promotion of slow pupils$^c$</td>
</tr>
<tr>
<td>21i15</td>
<td>preparation of written exams$^b$</td>
</tr>
<tr>
<td>21i16</td>
<td>advice on the assessment of student performance$^c$</td>
</tr>
<tr>
<td>21i17</td>
<td>joint promotion of high-performance students$^c$</td>
</tr>
</tbody>
</table>

---

$^a$ Dimension,” instruction.”

$^b$ Dimension,” project.”

$^c$ Dimension,” organization, performance and problems.”