Hier fancy Titel überlegen für Expertise-Paper

# Abstract

*Keywords:* classroom management, professional vision, eye-tracking, expertise differences, managing classroom disruptions

# Introduction

## Classroom Management

## Teachers’ Professional Knowledge

## Expertise Differences in Teachers’ Professional Vision

## Eye-Tracking to Assess Teachers’ Professional Vision

Fixation Number: This metric reflected the frequency of visual attention shifts, with higher fixation numbers indicating a more dynamic scanning behavior across the classroom environment (Grub et al., 2020).

Average Duration Fixation: This metric provided a measure of cognitive processing, with longer durations suggesting more time spent processing visual information (Negi & Mitra, 2020).

## Present Study

The present study aimed to investigate differences between experienced and inexperienced teachers in their gaze behavior during a micro-teaching unit, involving classroom disruptions their self-evaluations of competencies in classroom disruption prevention and management, their subjective assessments of how disruptive the events were, how confident they felt in dealing with them, and their strategic knowledge of classroom management. To address these objectives, data were analyzed from both in-service (experienced) and pre-service (inexperienced) teachers who participated in the laboratory-based study *Professional Vision of Novice and Expert Teachers* (ProVisioNET).

Participants individually attended a lab session and conducted a brief micro-teaching unit for a small “class” consisting of three trained actors who portrayed students and simulated typical classroom disruptions. During the micro-teaching unit, teachers’ gaze patterns were recorded using eye-tracking technology, while their evaluations of classroom management and strategic knowledge were assessed through self-report questionnaires, an interview, and a test.

The study addressed five research questions:

**Research question 1:** Which Areas of Interest (AOI) do teachers focus on during classroom interactions in a lab-based micro-teaching unit, lasting approximately 15 minutes, and where do they allocate their visual attention? Tentatively, teachers focused most frequently and for the longest durations on students, as this area was deemed the most relevant in the given classroom scenario. This exploration was descriptive and the percentage of fixation duration was used as the primary measure, as it offered a more intuitive visualization of attention distribution over time.

**Research question 2:** How do expertise differences manifest in gaze behavior? This question examined whether experienced teachers exhibit more efficient gaze behavior compared to inexperienced teachers, hypothesizing the following: Experienced teachers demonstrate more efficient gaze behavior, indicated by a lower Gaze Relational Index (GRI) for the entire micro-teaching unit than inexperienced teachers (**Hypothesis 2a**). Experienced teachers focus more frequently on students (higher mean number of fixations on AOI *Students*) than inexperienced teachers (**Hypothesis 2b**). In comparison to inexperienced teachers, experienced detect disruptions more quickly (shorter time to first fixations (TTFF) on the AOI *Disruptive Student*).

The second research goal was to examine global monitoring gaze behavior, represented by the Gaze Relational Index (GRI), calculated as the ratio of fixation number to fixation duration, for the entire micro-teaching unit and the AOI *Students*, as well as event-related gaze behavior to the AOI *Disruptive Person*. For the global monitoring behavior, experienced teachers were expected to demonstrate more frequent but shorter fixations, resulting in a smaller GRI than inexperienced teachers for the entire micro-teaching unit and the AOI *Students* (**Hypothesis 2a**). For the event-related gaze behavior, experienced teachers were expected to exhibit more frequent fixations and shorter fixation durations when attending to the AOI *Disruptive Person* compared to inexperienced teachers (**Hypothesis 2b**) and respond more quickly (shorter time to first fixation) to disruptions involving the AOI *Disruptive Person* compared to inexperienced teachers (**Hypothesis 2c**). Differences in event-related gaze behavior (fixation number, fixation duration, and TTFF) will be observed across disruption categories, with less salient disruptions (lack of eagerness to learn) receiving less visual attention compared to more salient categories such as verbal and physical disruptions (**Hypothesis 2d**).

1. The third research goal was to explore the relationship between gaze efficiency, as indicated by the GRI, and other classroom management characteristics. Teachers who felt more disrupted by classroom behaviors were expected to show higher GRI values, reflecting less effective gaze behavior (**Hypothesis 3a**). Teachers who felt more confident in dealing with the disruptions were expected to show lower GRI values, reflecting more effective gaze behavior (**Hypothesis 3b**). Furthermore, lower GRI values were hypothesized to correlate with higher self-assessed competencies in classroom disruption prevention and management (**Hypothesis 3c**). Additionally, lower GRI values were expected to correlate with greater strategic knowledge of classroom management (**Hypothesis 3d**).

# Method

## Participants

We recruited a total of 84 teachers from Germany (42 pre-service teachers and 42 in-service teachers) through personal contacts, email lists, and flyers. Pre-service teachers were required to be actively enrolled in a teacher education program and to have completed their first internship, while in-service teachers needed to have completed both the teacher training program (2. Staatsexamen) and to be currently working in the teaching profession. Data from two in-service teachers were excluded due to low-quality eye-tracking data, resulting in a final sample of 82 teachers, comprising 42 pre-service teachers and 40 in-service teachers.

The pre-service teachers (*n*= 29 women, *n*= 13 men) had a mean age of 22.80 years (*SD* = 1.90; range: 19–27). On average, they were in their 6.70 semesters (*SD* = 2.60; range: 3–11) and had an average of 9.60 hours (*SD* = 7.20; range: 1–36) of teaching experience through internships completed during their studies. Of the pre-service teachers, 21% were preparing to become primary school teachers, 60% were secondary school teachers, and 19% were special education teachers. Additionally, 88% were involved in extracurricular teaching activities, such as tutoring or homework supervision.

The in-service teachers (*n*= 24 women, *n*= 16 men) had a mean age of 39.10 years (*SD* = 10.60; range: 26–60) and an average of 11.60 years (*SD* = 11.30; range: 1–38) of teaching experience. Among these teachers, 10% taught at primary schools, 85% at secondary schools, and 5% at vocational schools. Furthermore, 52% were also involved in secondary teaching roles, such as university lecturers, main training supervisors for trainee teachers, or subject advisers.

The study adhered to ethical guidelines and received approval from the University’s Institutional Review Board. Participants were fully informed about the study’s objectives before testing. Their participation was voluntary, without incentives, and commenced only after written consent.

## Setting and Procedure

Participants individually attended the lab for approximately two hours, following a standardized procedure for which a seminar room was transformed into a classroom. Upon arrival, they were welcomed by the experimenter, introduced to the procedure, and asked to sign the data protection agreement. Participants were then fitted with eye-tracking glasses, adjusted for comfort and vision (up to +/- five diopters). After performing an initial one-point calibration of the glasses (for details of the calibration, see Eye-tracking apparatus and calibration), the experimenter activated and synchronized the recording devices (eye-tracking glasses, four cameras, and an audio recorder) using an auditory signal. This setup phase included a brief introductory game (“Name Juggling”) to acclimate participants to the eye-tracking equipment and the class, which took approximately 10-15 minutes.

After the initial setup, a second nine-point calibration was done in a separate room. As soon as the teacher re-entered the classroom, the micro-teaching unit started. Therefore, participants were asked to prepare a 15-minute micro-teaching unit on a topic and grade level of their choice. The only requirement was that the unit had to be an introductory micro-teaching unit, and had to consist of supervised individual work and/or frontal teaching. During the unit, three trained actors (playing students) performed scripted classroom disruptions, which occurred approximately every 1.5 minutes on a screen only visible to the “class” (e.g., chatting with a neighbor, heckling, looking at the phone; see Table A1 in the supplementary material for an overview and categorization of all events; and Figure B1 and B2 in the supplementary material for a depiction of the laboratory setting of the micro-teaching unit). The order of the disruptions and the performing students were fully balanced using Latin Squares. To capture teachers’ gaze patterns, the whole micro-teaching unit was recorded using eye-tracking glasses. The micro-teaching unit lasted about 15-20 minutes.

After the teaching session, participants repeated the nine-point calibration in the side room. During this time, the experimenter set up four letters A to C within the seminar room for a fixation task, which the participant performed after re-entering the room. Following this task, all recording devices were stopped, and participants filled out a brief computer-based questionnaire (~10-15 minutes) assessing sociodemographic data and a self-evaluation of their classroom management during the micro-teaching unit. A ten-minute break followed, concluding the first part of data collection.

In the study’s second phase, participants engaged in a Stimulated Recall Interview (SRI). They watched a video of their own teaching session, recorded through the eye-tracking glasses, while the experimenter paused the video at each classroom disruption. Participants answered five open-ended questions and three rating questions for each disruption, including self-evaluated disruption and confidence ratings (see Measures). The SRI lasted approximately 45-60 minutes. Finally, participants completed a digital Situational Judgment Test (SJT), assessing their strategic knowledge of classroom management. The questionnaire took approximately 15 minutes to complete, marking the end of the study.

## Eye-tracking apparatus and calibration

Teachers wore a binocular Tobii Pro Glasses 2 eye-tracker. The videos were recorded with a sampling rate of 50 Hz in a video resolution of 1920 x 1080 at 25 frames per second.

The evaluation of the calibration process followed the guidelines outlined in the Manual of Tobii AB (2024) and Onkhar et al. (2024) for assessing calibration quality.

## Measures

***Gaze Behavior***

Gaze behavior was analyzed using predefined Areas of Interest (AOIs; see Figure X) to examine how participants allocated their visual attention during the micro-teaching unit. Two types of AOIs were defined to structure the data meaningfully: global AOIs and event-based AOIs.

The global AOIs were used to assess gaze behavior over the entire video duration and included the following: the *Students*, representing gaze points focused on the group of three students; *Teacher Material*, capturing gaze points directed toward instructional materials such as the board, screen, or other teaching aids; *Student Desks*, representing gaze points on elements related to students’ desks, including name tags and student materials; and *Classroom/Others*, which encompassed gaze points directed toward other areas of the classroom that were not associated with the students or teacher.

In contrast, the event-based AOI focused on specific classroom events, such as disruptions caused by students. The AOI labeled *Disruptive Person* captured gaze behavior directed toward students performing the classroom disruptions.

#### Global Monitoring Gaze Behavior.

The global gaze monitoring behavior was assessed using three key eye-tracking metrics, recognized for their robustness and sensitivity to expertise in prior research (Grub et al., 2020). These metrics comprised fixation number, average fixation duration, and the Gaze Relational Index (GRI) as the ratio of these two measures (Gegenfurtner et al., 2020).

The **fixation number** was calculated as the total number of fixations recorded during the session, standardized as fixations per minute to account for varying session durations. Specifically, the total fixation count was summed across all global AOIs, including students, teacher material, student desks, and other classroom areas. This sum was then divided by the session duration in minutes to yield the fixation number per minute for each person and was then summarized for both groups (experienced and inexperienced teachers).

The **average fixation duration** was calculated to determine the mean time participants spent focusing on individual elements. This was computed by dividing the total duration of all fixations, measured in milliseconds, by the total number of fixations recorded during the session. Fixation data for these calculations were aggregated across all global AOIs, offering a comprehensive overview of participants’ monitoring behavior throughout the micro-teaching unit for each person.

The **GRI** was calculated as a composite metric to evaluate the efficiency of participants’ scanning behavior during the micro-teaching unit. The GRI was derived by dividing the average fixation duration (in milliseconds) by the fixation number per minute. This index provided a standardized measure of visual scanning efficiency, with smaller GRI values indicating a combination of shorter fixation durations and higher fixation frequencies, which is typically associated with more dynamic and efficient scanning behavior (Gegenfurtner et al., 2020).

#### Event-related Gaze Behavior.

Event-related gaze behavior allowed for a nuanced analysis of participants’ visual attention allocation in response to disruptions, providing insights into differences in attentional strategies between experienced and inexperienced teachers. This behavior was analyzed by focusing on specific classroom disruptions during the micro-teaching unit, categorized into three types: verbal disruptions (e.g., chatting with a neighbor, heckling, whispering), physical disruptions (e.g., clicking a pen, drumming with hands, snapping fingers), and indicators of lack of eagerness to learn (e.g., looking at a phone, resting the head on a desk, drawing). These disruptions were captured within the event-based AOI labeled *Disruptive Person*, which represented visual fixations directed toward students engaging in disruptive or off-task behaviors.

Three key metrics were calculated to assess gaze behavior within this AOI:

First, the **number of fixations per second** quantified how frequently participants fixated on the *Disruptive Person* during the disruption events. The total number of fixations within the AOI was divided by the actual duration of the disruption in seconds to calculate the fixation rate.

Second, the **average fixation duration** represented the mean time, in milliseconds, that participants spent fixating on the AOI *Disruptive Person* during each fixation. It was calculated by dividing the total duration of fixations within the AOI by the total number of fixations directed at the AOI.

Third, the **time to first fixation** (TTFF) was used to measure the latency, in seconds, for participants to direct their first fixation to the AOI after the onset of the disruption. Therefore, values were extracted from the eye-tracking data by identifying the timestamp when the first fixation on the AOI *Disruptive Student* occurred relative to the onset of the disruptive event. Data were filtered to include only valid fixation times, excluding instances where the TTFF was zero[[1]](#footnote-2) or exceeded 30 seconds.[[2]](#footnote-3) These raw TTFF values were then converted from milliseconds to seconds for interpretability. To calculate participant-level averages, the TTFF values were aggregated across all disruption events per participant. This yielded an average TTFF in seconds for each participant, reflecting their overall responsiveness to classroom disruptions.

### Teaching Experience

Participants’ teaching experience was assessed as part of their sociodemographic data, with the duration of their work experience reported in years.

### Self-Evaluations of Competencies in Classroom Disruption Prevention and Management

After the micro-teaching unit, teachers answered a questionnaire using five items from a validated questionnaire (Helmke et al., 2013) and nine self-developed items derived from the research literature (Kiel et al., 2013; Kounin, 2006; Marzano, 2007) to assess teachers’ self-evaluated competencies in classroom disruption prevention and management. The questionnaire was a 4-point Likert scale (1 = strongly disagree; 4 = strongly agree).

### Subjective Disruption and Confidence Ratings of the Classroom Events

The subjective disruption and confidence ratings were assessed during the SRI on an 11-point rating scale, ranging from 0 (not at all disrupting/confident) to 10 (extremely disrupting/confident). Ratings were averaged across the nine classroom events for each participant to capture a general sense of how disruptive the events were during the micro-teaching unit and how confident participants felt in handling the classroom events.

### Strategic Knowledge of Classroom Management

Teachers’ strategic knowledge of classroom management was assessed using a Situational Judgment Test (SJT; Gold & Holodynski, 2015) via an online questionnaire on SoSci Survey. Participants graded five to six action alternatives for twelve teaching scenarios in which classroom disruptions were discussed on a six-point Likert scale (grade 1 = “very good” to grade 6 = “unsatisfactory”). As the SJT was originally designed for primary schools, adjustments were made in order to be able to use the SJT for all types of schools in the *ProVisioNET* study. Due to their general applicability, all twelve scenarios and answer options were adopted and only the names of the class levels were removed from the questions - except for scenario 6, where this information was essential. Only fully completed questionnaires were included in the present study.

## Data analysis

The eye-tracking data were analyzed using the software Tobii Pro Lab Analyzer (Version 1.241.54542). A fixation filter was applied, with a threshold set at 30°/sec to identify fixations, as this default fixation filter is recommended for mobile eye-tracking data in the Tobii Lab Analyzer Software (Tobii AB, 2024).

The data were analyzed using R (RStudio Team, 2020, Version 2024.12.0) and IBM SPSS Statistics (Version 29). Graphics were created using ggplot2 (Wickham, 2016).

To test **Hypothesis 1**, we analyzed the percentage of fixations and fixation durations directed at each AOI during the micro-teaching unit. The AOIs included *Students, Disruptive Student, Teacher Material, Student Desks*, and *Classroom/Others*. For fixation percentages, we calculated the proportion of fixations on each AOI relative to the total number of fixations, excluding fixations on the *Disruptive Student*. For fixation durations, we calculated the proportion of fixation time on each AOI relative to the total fixation duration, again excluding durations associated with the *Disruptive Student*. A separate percentage for the *Disruptive Student* was computed using denominators that included the fixations and durations on this AOI to ensure proportional comparisons across all categories. For each AOI, mean percentages of fixations and fixation durations were calculated separately for experienced and inexperienced teachers. These values were visualized using bar graphs, with the AOIs as categories and percentages as dependent variables, allowing direct comparisons between groups. To determine statistical differences between the two groups, independent-sample *t*-tests were conducted for each AOI, both for fixation percentages and fixation durations. Effect sizes were reported using Cohen’s *d*, providing information on the magnitude of group differences.

To test **Hypothesis 2**, which examined global monitoring gaze behavior and event-related gaze behavior, we conducted a series of analyses. For global monitoring gaze behavior, represented by the Gaze Relational Index (GRI), we analyzed data from the entire micro-teaching unit and the AOI *Students*. The GRI was calculated as the ratio of fixation number to fixation duration, with a smaller GRI reflecting more frequent but shorter fixations (**Hypothesis 2a)**. Independent-sample *t*-tests were used to compare the group means of fixation number, fixation duration, and GRI between experienced and inexperienced teachers. Cohen’s *d* was calculated to determine the effect size, indicating the magnitude of observed differences. To test **Hypothesis 2b**, which proposed that experienced teachers would exhibit more frequent fixations and shorter fixation durations when attending to the AOI *Disruptive Student* compared to inexperienced teachers, we analyzed the total fixation number and fixation duration directed toward this AOI. Independent-sample *t*-tests were used to compare these parameters between groups. Cohen’s *d* was calculated to quantify the magnitude of the differences. To test **Hypothesis 2c**, which posited that experienced teachers would respond more quickly (shorter time to first fixation) to disruptions involving the AOI *Disruptive Student* compared to inexperienced teachers, we analyzed the TTFF across all disruption events. Independent-sample *t*-tests were conducted to compare the TTFF between experienced and inexperienced teachers. Cohen’s *d* was again calculated to assess the effect size of the group differences. To test **Hypothesis 2d**, which stated that differences in gaze behavior (fixation number, fixation duration, and TTFF) between experienced and inexperienced teachers would vary across disruption categories (verbal, physical, and lack of eagerness to learn), we performed a two-way repeated-measures Analysis of Variance. The analysis examined the main effects of group (experienced vs. inexperienced), disruption category, and their interaction for each parameter. Generalized eta-squared (*η²*) was calculated to quantify the effect sizes for each factor. Pairwise comparisons were conducted to (1) identify group differences (experienced vs. inexperienced teachers) within each disruption category and (2) examine differences between the disruption categories (e.g., less salient disruptions, such as lack of eagerness to learn, compared to more salient categories like verbal or physical disruptions). Cohen’s *d* was calculated for pairwise comparisons to quantify the magnitude of differences between groups and across disruption categories.

To address the third research goal, we analyzed the relationship between gaze efficiency (GRI) and classroom management characteristics using Pearson correlation coefficients. For each hypothesis, GRI values were correlated with classroom management measures: ratings of how disturbing the disruption was (**Hypothesis 3a**), ratings of confidence in managing disruptions (**Hypothesis 3b**), self-assessed competencies in disruption prevention and management (**Hypothesis 3c**), and strategic knowledge of classroom management (**Hypothesis 3d**). Significance was assessed with *p*-values.

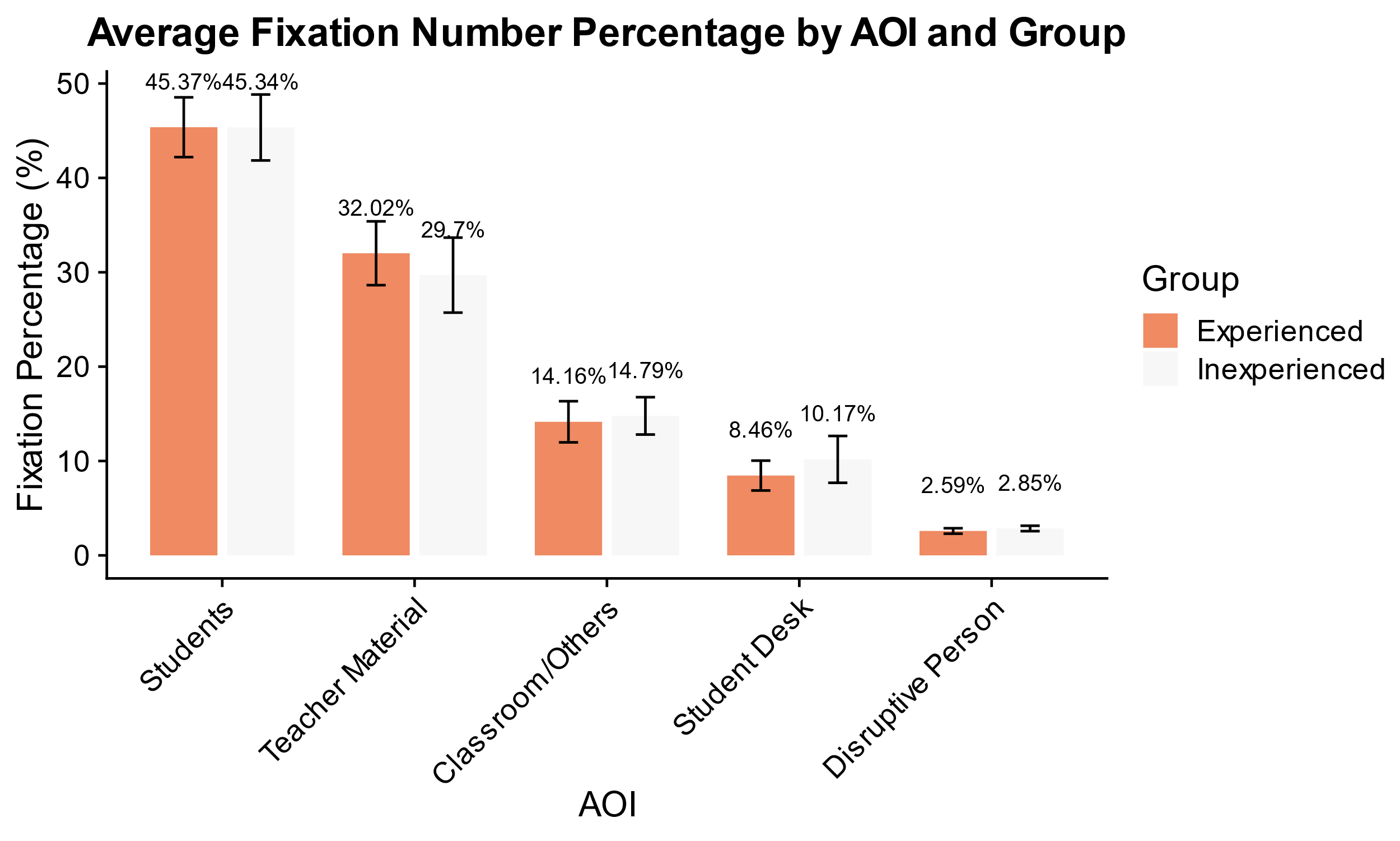
# Results

## Gaze Behavior Across Areas of Interest

To investigate teachers’ gaze allocations, we calculated the percentage of fixation numbers and fixation durations directed at each AOI during the micro-teaching unit. The results for experienced and inexperienced teachers are displayed in Figures 1 and Figure 2.

**Figure 1**

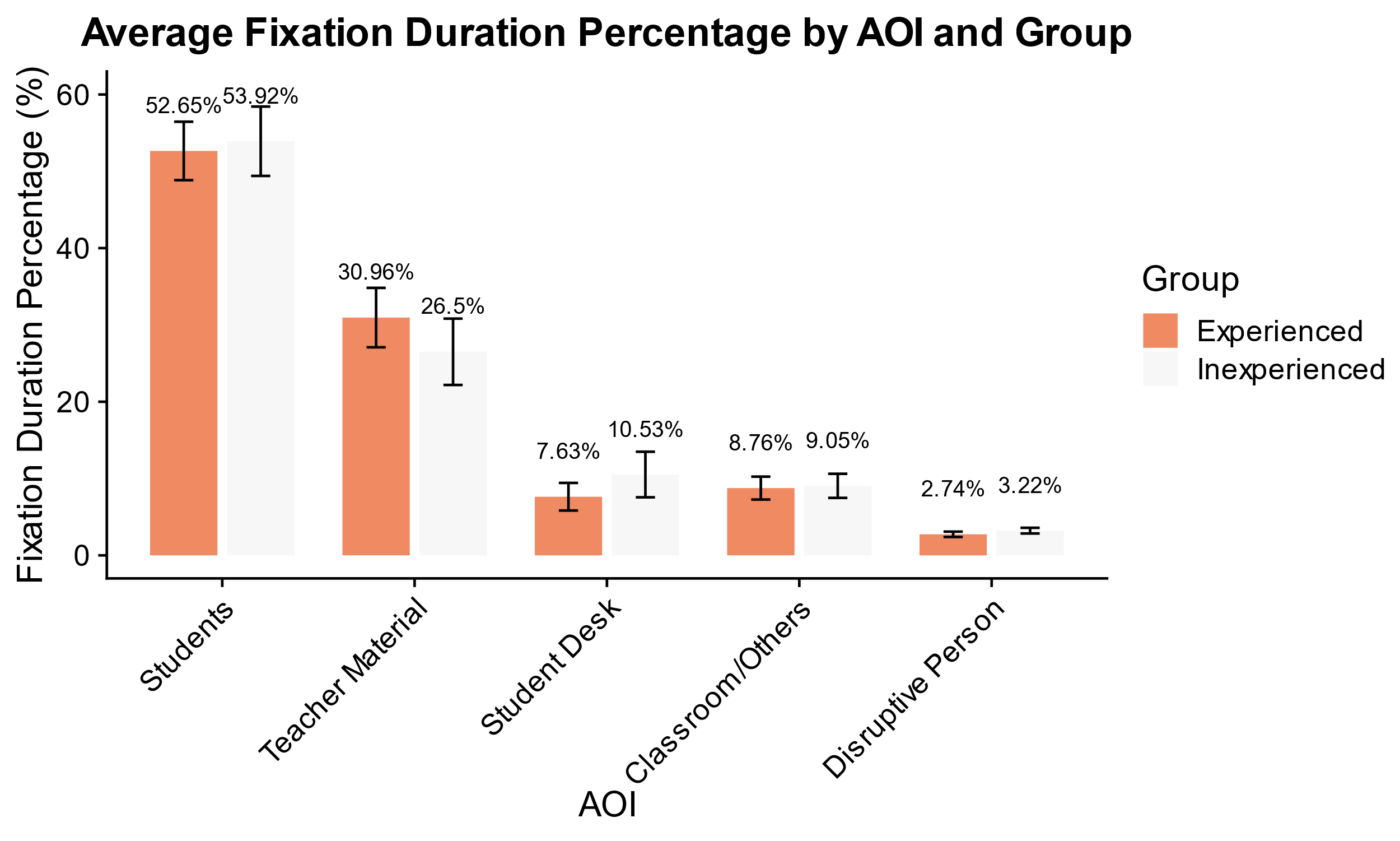
*Average Fixation Number Percentages by Area of Interest (AOI) and Teacher Experience Group for the Entire Micro-Teaching Unit with 95% Confidence Intervals*

**

*Note.* The bar chart illustrates the average fixation number percentage directed at five AOIs (Students, Teacher Material, Classroom/Others, Student Desk, and Disruptive Person) during the micro-teaching unit. Results are presented separately for experienced and inexperienced teachers, with error bars indicating 95% confidence intervals.

**Figure 2**

*Average Fixation Duration Percentages by Area of Interest (AOI) and Teacher Experience Group for the Entire Micro-Teaching Unit with 95% Confidence Intervals*

******

*Note.* The bar chart displays the average fixation duration percentage across five AOIs (Students, Teacher Material, Student Desk, Classroom/Others, and Disruptive Person) during the micro-teaching unit. Data are presented separately for experienced and inexperienced teachers, with error bars indicating 95% confidence intervals.

As predicted in **Hypothesis 1**, both groups exhibited the highest percentages of fixation number and average fixation duration directed toward the AOI *Students*, followed by the AOI *Teacher Material*. The AOI *Disruptive Person* received the lowest fixation percentages. For both measures, no statistically significant differences were observed between experienced and inexperienced teachers across any AOI (see Appendix A, Table A1).

**Table A1**

t*-Test Results and Effect Sizes for Fixation Number Percentages (FNP) and Fixation Duration Percentages (FDP) Across AOIs Between Experienced and Inexperienced Teachers*

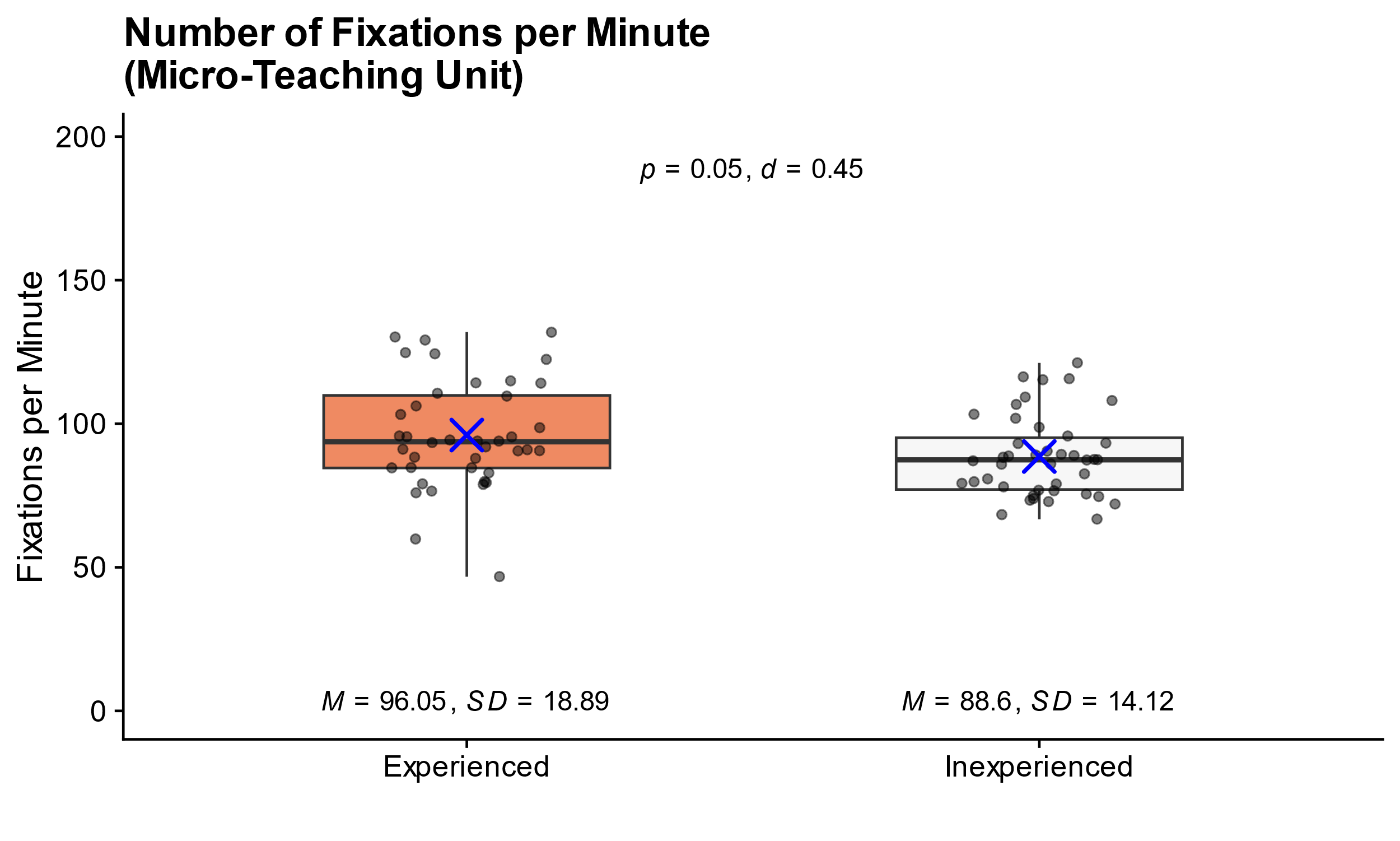
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Measure | AOI | *t-*value | df | *p*-value | *M* Difference | Cohen's *d* | 95% CI |
| FNP a | Students | 0.01 | 80 | 0.99 | 0.03 | 0.00 | [–4.63, 4.69] |
| FNP | Disruptive Person | –1.32 | 80 | 0.19 | –0.27 | –0.29 | [–0.67, 0.14] |
| FNP | Teacher Material | 0.89 | 80 | 0.37 | 2.32 | 0.20 | [–2.84, 7.48] |
| FNP | Student Desk | –1.17 | 80 | 0.25 | –1.72 | –0.26 | [–4.65, 1.21] |
| FNP | Classroom/Others | –0.43 | 80 | 0.67 | –0.63 | –0.10 | [–3.53, 2.27] |
| FDP b | Students | –0.43 | 80 | 0.67 | –1.26 | –0.10 | [–7.1, 4.58] |
| FDP | Disruptive Person | –1.90 | 80 | 0.061 | –0.48 | –0.42 | [–0.98, 0.02] |
| FDP | Teacher Material | 1.55 | 80 | 0.13 | 4.45 | 0.34 | [–1.28, 10.18] |
| FDP | Student Desk | –1.67 | 80 | 0.1 | –2.89 | –0.37 | [–6.35, 0.56] |
| FDP | Classroom/Others | –0.28 | 80 | 0.78 | –0.30 | –0.06 | [–2.43, 1.84] |
| *Note*. This figure displays the results of *t*-tests and effect sizes (Cohen’s *d*) for fixation number percentages (FNP) and fixation duration percentages (FDP) across Areas of Interest (AOIs) between experienced and inexperienced teachers. AOIs included Students, Disruptive Person, Teacher Material, Student Desk, and Classroom/Others. Positive mean differences (*M* Difference) indicate higher percentages for experienced teachers. Confidence intervals (95%) for the mean differences are presented in brackets.  *p*-value < .05 is considered statistically significant.  a FNP = Fixation Number Percentages  b FDP = Fixation Duration Percentages | | | | | | | |

## Global Monitoring and Event-related Gaze Behavior

To examine global monitoring behavior, we tested whether experienced teachers exhibit more frequent but shorter fixations, resulting in a smaller GRI than inexperienced teachers, first for the entire micro-teaching unit and second for the AOI *Students* (**Hypotheses 2a**).

Figure 3 shows that experienced teachers had a more fixations per minute (*M* = 96.05, *SD* = 18.89) compared to inexperienced teachers (*M* = 88.6, *SD* = 14.12) during the entire micro-teaching unit. Results from the *t*-test revealed a significant difference between the two groups, *t*(80) = 2.03, *p* = .05, *d* = 0.45 (small effect).

**Figure 3**

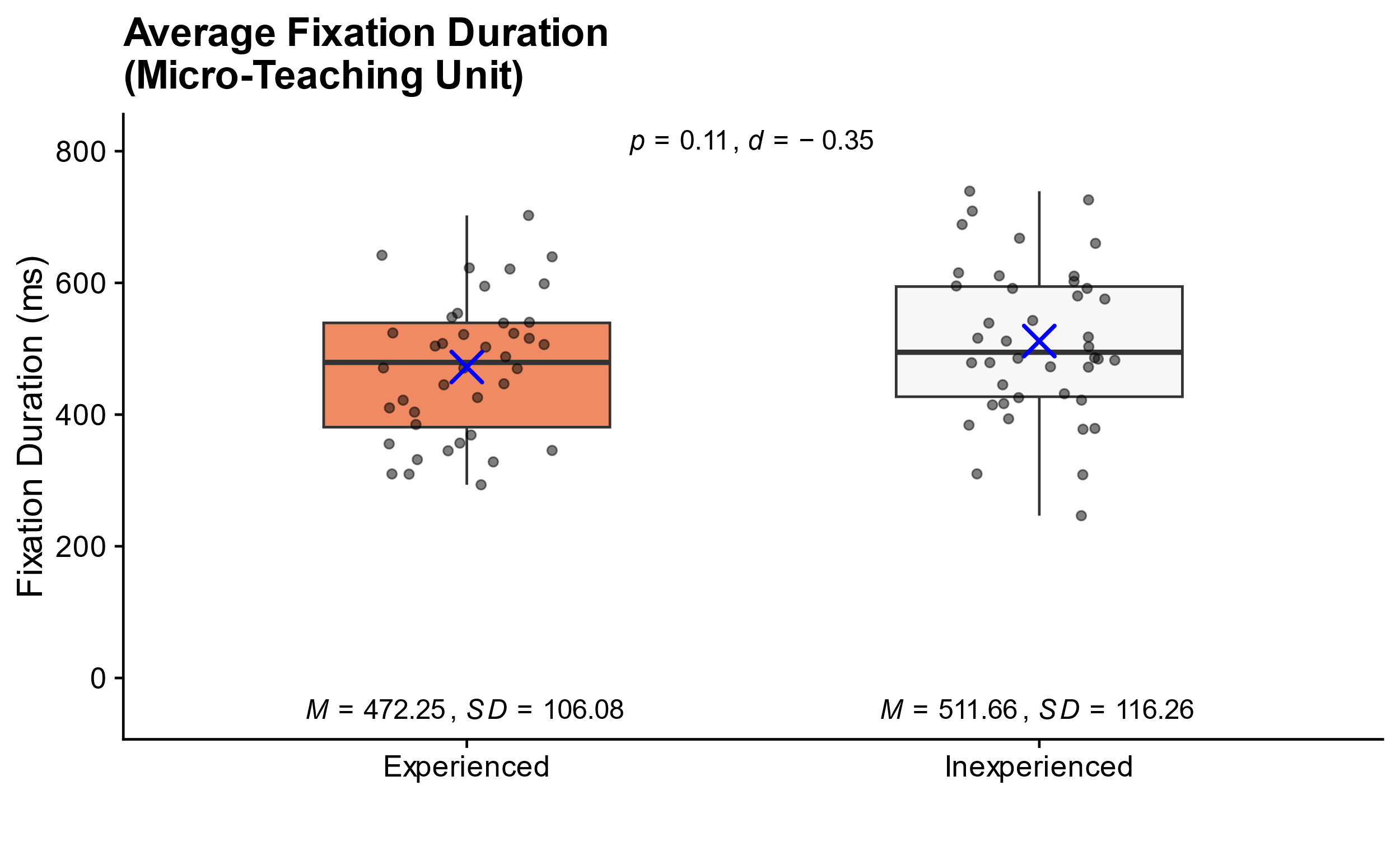
*Number of Fixations per Minute in the Micro-Teaching Unit for Experienced and Inexperienced Teacher*

*Note*. Number of fixations per minute in the micro-teaching unit for 40 experienced and 42 inexperienced participants. The boxplot displays the median (bold line), interquartile range (box spanning the 25th to 75th percentiles), whiskers extending to 1.5 times the interquartile range, and individual data points (dots). The blue "X" indicates the group mean, while the black error bars represent the standard deviation.

As shown in Figure 4, experienced teachers (*M* = 472.25 ms, *SD* = 106.08 ms) exhibited shorter fixation durations than inexperienced participants (*M* = 511.68 ms, *SD* = 116.26 ms). However, this difference was not statistically significant, *t*(80) = –1.60, *p* = .11, *d* = –0.35 (small effect).

**Figure 4**

*Average Fixation Duration in Milliseconds in the Micro-Teaching Unit for Experienced and Inexperienced Teachers*

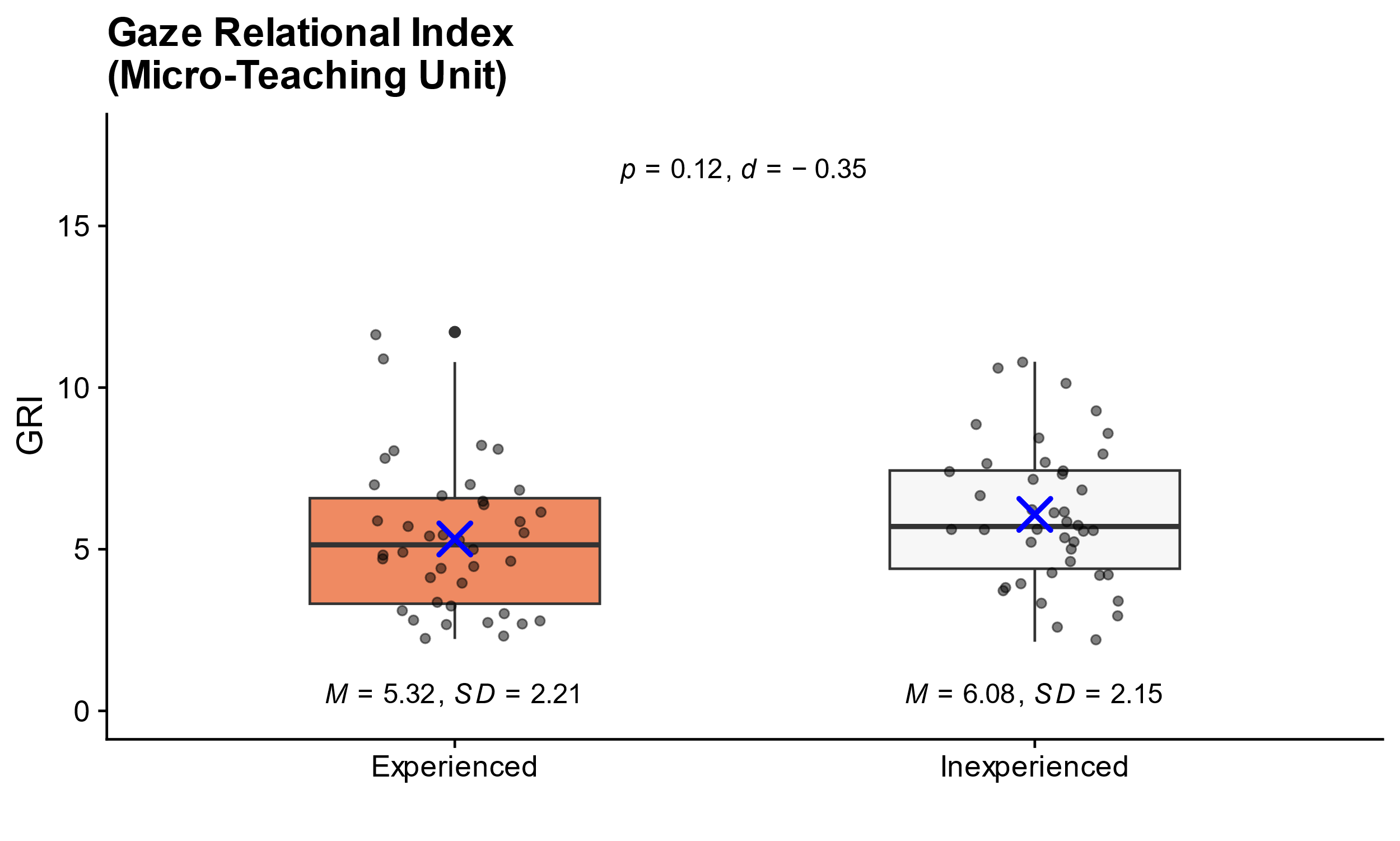


*Note*. Average fixation duration in milliseconds in the micro-teaching unit for 40 experienced and 42 inexperienced participants. The boxplot displays the median (bold line), interquartile range (box spanning the 25th to 75th percentiles), whiskers extending to 1.5 times the interquartile range, and individual data points (dots). The blue "X" indicates the group mean, while the black error bars represent the standard deviation.

As depicted in Figure 5, inexperienced teachers (*M* = 6.08, *SD* = 2.15) had a slightly higher GRI values for the entire micro-teaching unit compared to experienced teachers (*M* = 5.32, *SD* = 2.21). However, this difference was not statistically significant, *t*(80) = −1.58, *p* = .12, *d* = −0.35 (small effect).

**Figure 5**

*Gaze Relational Index (GRI) in the Micro-Teaching Unit for Experienced and Inexperienced Teachers*

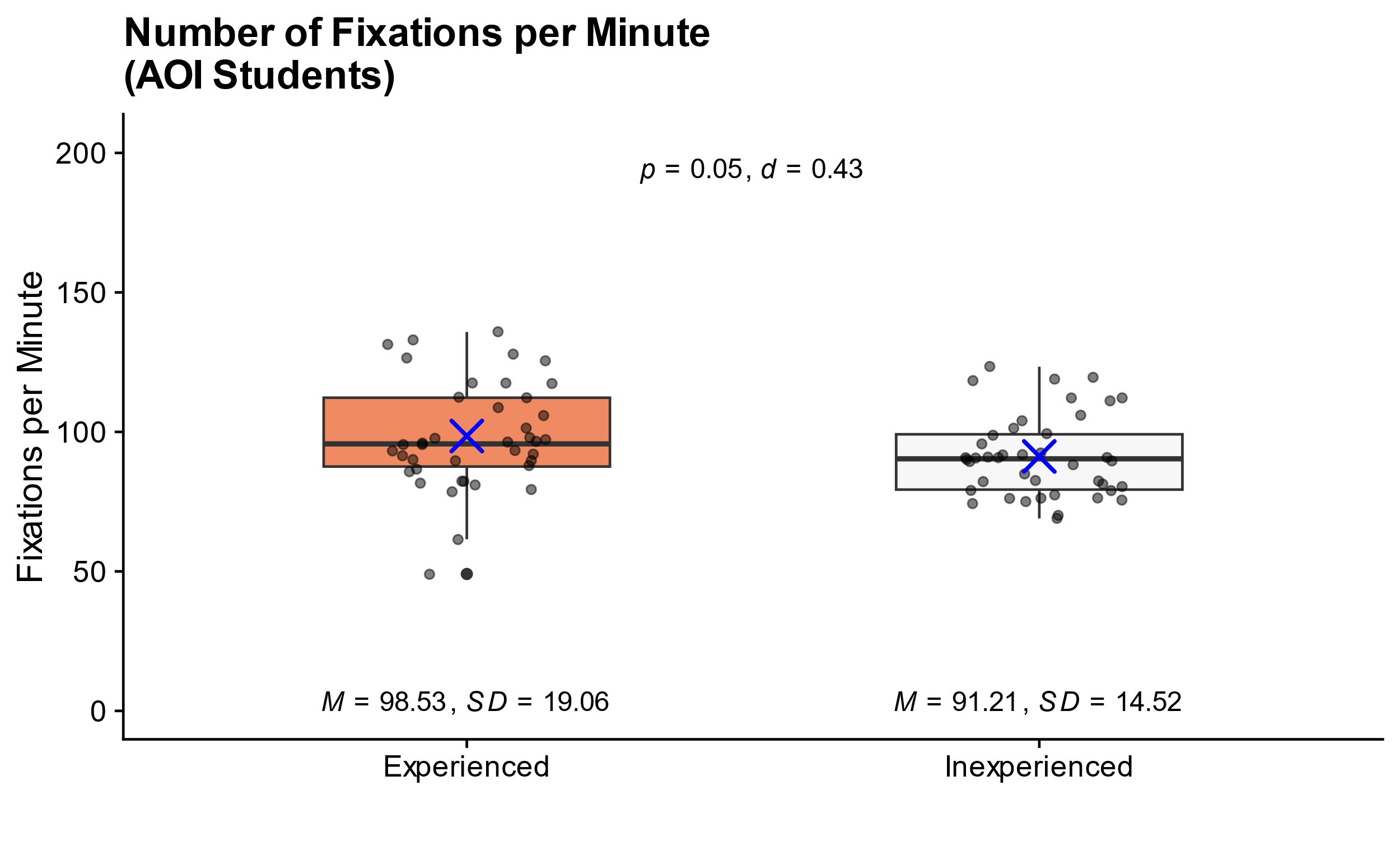


*Note.* The Gaze Relational Index (GRI) is calculated as the ratio of the average fixation duration to the number of fixations per minute. The boxplots represent the interquartile range and the median (bold line), with whiskers extending to 1.5 times the interquartile range. Individual data points are displayed as dots and the blue “X” indicates the group mean.

As shown in Figure 6, experienced teachers (*M* = 98.53, *SD* = 19.06) directed their gaze toward AOI *Students* more frequently than inexperienced teachers (*M* = 91.21, *SD* = 14.52). This difference was statistically significant, *t*(80) = 1.96, *p* = .05, *d* = 0.43 (small effect size).

**Figure 6**

*Number of Fixations per Minute Directed at AOI Students by Experienced and Inexperienced Teachers*

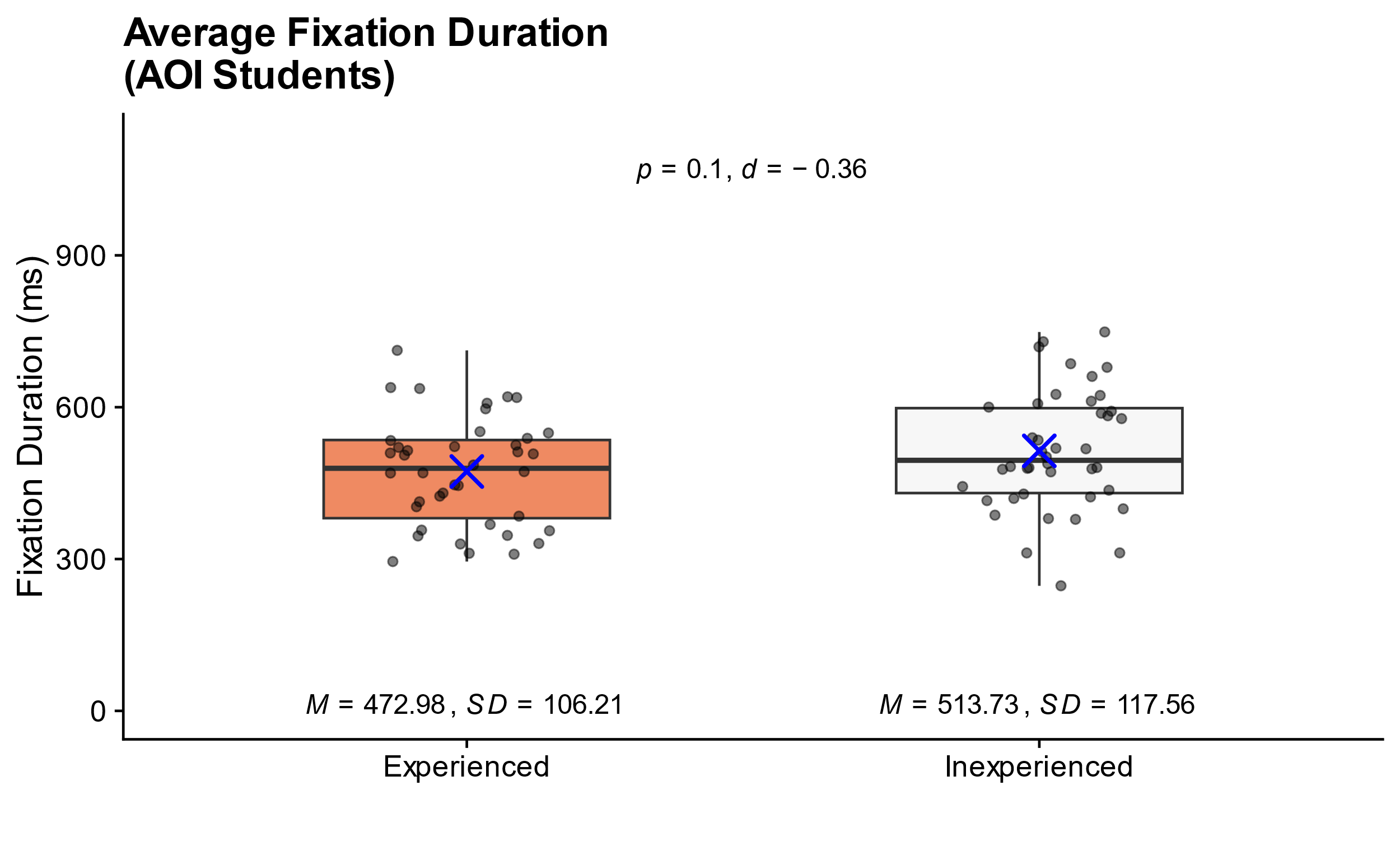


*Note.* The figure depicts the number of fixations per minute directed at AOI *Students* by experienced and inexperienced teachers. The boxplots represent the interquartile range and the median (bold line), with whiskers extending to 1.5 times the interquartile range. Individual data points are displayed as dots and the blue “X” indicates the group mean.

Regarding the average duration of fixations on the AOI *Students*, experienced teachers (*M* = 472.98 ms, *SD* = 106.21 ms) exhibited shorter fixation durations compared to inexperienced teachers (*M* = 513.73 ms, *SD* = 117.56 ms) as shown in Figure 7. However, this difference was not statistically significant, *t*(80) = −1.64, *p* = .10, *d* = −0.36 (small effect size).

**Figure 7**

*Average Fixation Duration Directed at AOI Students by Experienced and Inexperienced Teachers.*



*Note.* The figure depicts the average fixation duration (in milliseconds) directed at AOI *Students* by experienced and inexperienced teachers. The boxplots represent the interquartile range and the median (bold line), with whiskers extending to 1.5 times the interquartile range. Individual data points are displayed as dots and the blue “X” indicates the group mean.

To examine evet-related gaze behavior, we tested whether exhibit more frequent fixations and shorter fixation durations when attending to the AOI *Disruptive Person* compared to inexperienced teachers (**Hypotheses 2b**).

# Discussion

## Key Findings

Our study investigated …

Overall, our findings indicate that …

Our findings are consistent with prior research that illustrates the …

## Limitations and future directions

While the laboratory setting of the study allowed for a controlled implementation of stressors and high internal validity, it was not an authentic classroom environment, raising questions about its external validity. Most importantly, the teacher and their students did not have a shared history, and only a very thin basis for establishing a positive teacher-student relationship, which is a core characteristic of effective classroom management (Beaty-O’Ferrall et al., 2010; Rüedi, 2014).

In addition, the micro-teaching unit was only about 15 minutes long, and thus much shorter than a regular school lesson, providing less opportunities for experienced teachers to build up an engaging lesson.

Finally, the onset of disruptive student behavior was scripted, following an experimental time schedule, which was not affected by the behavior of the teacher. Thus, the setting may have masked effects of teaching experience by providing too little opportunities of experienced teachers to demonstrate their true classroom management skills.

In subsequent studies, it would therefore be insightful to …

## Conclusion

This study investigated …

In summary, our study contributes to the understanding …

# References

Beaty-O’Ferrall, M. E., Green, A., & Hanna, F. (2010). Classroom Management Strategies for Difficult Students: Promoting Change through Relationships. *Middle School Journal (J1)*, *41*(4), 4–11.

Gegenfurtner, A., Boucheix, J.-M., Gruber, H., Lehtinen, E., & Lowe, R. K. (2020). *The Gaze Relational Index as a Measure of Visual Expertise*. *3*.

Gold, B., & Holodynski, M. (2015). Development and Construct Validation of a Situational Judgment Test of Strategic Knowledge of Classroom Management in Elementary Schools. *Educational Assessment*, *20*(3), 226–248. https://doi.org/10.1080/10627197.2015.1062087

Grub, A.-S., Biermann, A., & Brünken, R. (2020). *Process-based measurement of professional vision of (prospective) teachers in the field of classroom management. A systematic review*. https://doi.org/10.25656/01:21187

Helmke, A., Schrader, F.-W., Helmke, T., Lenske, G., Pham, G., Praetorius, A.-K., & Ade-Thurow, M. (2013). *Basisfragebogen EMU - Evidenzbasierte Methoden der Unterrichtsentwicklung*. Unterrichtsdiagnostik. http://www.unterrichtsdiagnostik.info/downloads/fragebogen/

Kiel, E., Frey, A., & Weiß, S. (2013). *Trainingsbuch Klassenführung*. Verlag Julius Klinkhardt.

Kounin, J. S. (2006). *Techniken der Klassenführung*. Waxmann Verlag.

Lohmann, G., & Meyer, H. (2003). *Mit Sch"ulern klarkommen: Professioneller Umgang mit Unterrichtsst"orungen und Disziplinkonflikten*. Cornelsen-Scriptor.

Marzano, R. J. (2007). *The Art and Science of Teaching: A Comprehensive Framework for Effective Instruction*. ASCD.

Onkhar, V., Dodou, D., & de Winter, J. C. F. (2024). Evaluating the Tobii Pro Glasses 2 and 3 in static and dynamic conditions. *Behavior Research Methods*, *56*(5), 4221–4238. https://doi.org/10.3758/s13428-023-02173-7

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC.

Rüedi, J. (2014). Zur Bedeutung positive Beziehungen für die Klassenführung und den Umgang mit Unterrichtsstörungen. *Beziehungen in Schule Und Unterricht. Teil*, *3*, 105–126.

Tobii AB. (2024). *Tobii Pro Lab User Manual v 24.21*. https://go.tobii.com/tobii\_pro\_lab\_user\_manual

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://doi.org/10.1007/978-0-387-98141-3

# Appendix

# Appendix A

**Table A1**

*Classification of Nine Typical Classroom Disruptions According to Lohmann & Meyer (2003) Performed in The Micro-Teaching Unit by Actors*

|  |  |  |
| --- | --- | --- |
| Verbal disruptions | Physical disruptions | Lack of eagerness to learn |
| Heckling | Clicking pen | Looking at phone |
| Chatting | Snipping hands | Drawing |
| Whispering | Drumming hands | Head on table |

*Note.* Disruptions were classified based on the typology provided by Lohmann & Meyer (2003)*.* Categories include verbal, physical, and disengagement-related behaviors performed during the micro-teaching unit. The order of the performing actors and the disruptions was fully balanced using Latin squares.

# Appendix B

# Laboratory Setting of The Study

**Figure B1**

*Laboratory Setting of The Micro-Teaching Unit. Ein Bild, das Mobiliar, Stuhl, Kleidung, Schuhwerk enthält.

Automatisch generierte Beschreibung*

*Note*. The setting included three actors as the class (left) and a teacher (participant, right).

**Figure B2**

*Laboratory Setting of The Interview.*

Ein Bild, das Mobiliar, Zeichnung, Entwurf, Tisch enthält.

Automatisch generierte Beschreibung

*Note*. The experimenter and participant watched the previously taught micro-teaching unit on video.

# Appendix C

1. As our focus was on gaze behavior directed toward the disruptive student, instances where the gaze was already fixated on the disruptive student, when the event started, were excluded from the analysis. [↑](#footnote-ref-2)
2. Fixation times beyond 30 seconds were excluded, as they exceeded the duration of the scripted disruptions. [↑](#footnote-ref-3)