Contemporary Educational Psychology 64 (2021) 101928

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

A micro-perspective on students' behavioral engagement in the context of teachers' instructional support during seatwork: Sources of variability and the role of teacher adaptive support☆

Marije van Braak[a,1,*], Janneke van de Pol[a], Astrid M.G. Poorthuis[b], Tim Mainhard[a]

[a] Utrecht University, Department of Education, the Netherlands
[b] Utrecht University, Department of Psychology, the Netherlands

A R T I C L E I N F O

Keywords:
Behavioral engagement
On-task behavior
Adaptive support
Cross-classified multilevel modeling
Classroom observation

A B S T R A C T

Despite increased acknowledgement of the significance of situational factors to engagement, engagement is traditionally seen as a student characteristic. In this study, we investigated to what extent variation in observational measures of behavioral student engagement during seatwork is due to students versus teachers, teacher-student dyads, or situational (i.e., variation in time) effects. Additionally, we studied whether adaptive support during teacher-student interactions contributes to student engagement following that interaction. A cross-classified multilevel analysis of 324 video- and audio-recorded secondary school teacher-student interactions showed that situational factors and error were the strongest source of variability in student engagement. Other than expected, teacher-related and dyad-related variability were limited. Although behavioral engagement after teacher-student interaction was significantly higher than pre-interaction student engagement, higher post-interaction levels were not related to the level of adaptive support provided during the teacher-student interaction. These findings imply that situational factors deserve more attention in engagement research, for example by investigating what situational factors contribute to engagement. For teachers, the time-variant, situation-dependent nature of engagement opens up valuable opportunities to actively design optimal learning situations.

1. Introduction

Historically, school engagement has mainly been studied as a student characteristic (Pöysä et al., 2018), indicating that some students tend to be engaged while others are generally disengaged in class. Current research increasingly acknowledges the malleability and responsiveness of engagement to variation in learning contexts (Christenson, Reschly, & Wylie, 2012; Fredricks & McColskey, 2012; Pöysä et al., 2018). The current study adds to this focus by examining student engagement in the context of teachers' instructional support at a micro, moment-to-moment level.

In past studies differences in engagement between students may have been overestimated in two ways. First, if a student's engagement is assessed at one point in time only, potential *situational* fluctuations in a student's engagement are confounded with stable differences in engagement between individual students. Second, if a student's engagement is assessed in classes of one teacher only, it is not possible to distinguish between the general or habitual level of a student's engagement and the possible attunement between a specific teacher and a specific student. Omitting these so-called *relationship* effects (Mainhard, Oudman, Hornstra, Bosker, & Goetz, 2018) may lead to overestimating differences between students. We aimed at estimating these possible sources and predictors of variability in engagement.

Theoretically, knowledge about sources of variation in engagement is important for both conceptual clarity and operationalization (Fredricks & McColskey, 2012). Understanding the sources of variability in engagement is important for the selection of predictors at the respective levels in predictive statistical models. For example, if variation in engagement would be mainly situational, predictors at that level (e.g., how adaptively a teacher responds to a students' needs in a specific situation) are most probable to explain variance in engagement. Practically, knowledge about sources and predictors of variation in engagement may point to the appropriate levels for interventions aimed at enhancing engagement in class.

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author at: Department of General Practice, Erasmus Medical Centre, Postbus 2040, Dr. Molewaterplein 40, 3015 GD Rotterdam, the Netherlands.
E-mail address: m.vanbraak@erasmusmc.nl (M. van Braak).
[1] Present address: Erasmus Medical Center, Department of General Practice, Rotterdam, the Netherlands.

https://doi.org/10.1016/j.cedpsych.2020.101928

Available online 5 December 2020

As we were specifically interested in the importance of situational factors for student engagement, we explored teachers' adaptive instructional support during seatwork as a predictor of students' behavioral engagement. Teachers' adaptive support in interaction with students is of specific relevance to students' engagement at several levels. Some teachers may be generally better at providing support and engage students through their interactions than others (Tropper, Leiss, & Hänze, 2015) whereas students may generally differ in their responsiveness to these interactions. At the same time, teacher-student interaction is susceptible to the specific situation and this may be crucial for how well a student is able to carry on with working on school assignments.

Thus, we take a unique micro-perspective on engagement variability by describing and predicting behavioral engagement in the context of instructional support during *seatwork*. Seatwork is a classroom setting that is particularly prevalent in contemporary secondary education (Hiebert, 2003) and requires students to work independently and stay engaged with and focused on their work. Student engagement is therefore particularly important in this context.

2. Behavioral engagement

Behavioral engagement forms a major link between learning opportunities offered and actual learning (Appleton, Christenson, Kim, & Reschly, 2006; Fredricks, Blumenfeld, & Paris, 2004; Skinner, Kindermann, & Furrer, 2009). The concept centers around student participation (Appleton, Christenson, & Furlong, 2008; Reschly & Christenson, 2006; Shernoff et al., 2016), but has also been conceptualized as effort (Appleton et al., 2008; Shernoff et al., 2016; Skinner et al., 2009), attention (Finn & Zimmer, 2012; Guo, Sun, Breit-Smith, Morrison, & Connor, 2015; Skinner & Pitzer, 2012), attendance (Mahatmya, Lohman, Matjasko, & Farb, 2012; Reschly & Christenson, 2006; Shernoff et al., 2016), action initiation (Guo et al., 2015; Skinner & Pitzer, 2012), and time on task (Finn & Zimmer, 2012; Guo et al., 2015).

Behavioral engagement is presumed to be malleable (Appleton et al., 2008; Fredricks et al., 2004), "responsive to variation in the context" (Fredricks & McColskey, 2012, p. 765), and inextricably linked to the individual's environment (Christenson et al., 2012). As such, behavioral engagement can be seen as students' situation-dependent involvement in task-related activities. Although it is the student that displays engagement, this view on engagement highlights the interactional nature of engagement: being engaged as a "constantly fluctuating state sensitive to situational factors" (Pöysä et al., 2018, p. 65). Fellow students, teachers, tasks, and environmental features such as time of the day and classroom atmosphere are all part of this potentially influencing context (Eccles & Wang, 2012; Fredricks et al., 2004; Pöysä et al., 2018; Skinner et al., 2009).

3. Sources of variability in behavioral engagement

Earlier variance decompositions of engagement in general and specifically of behavioral engagement in classroom settings provide evidence for the relevance of a variety of potential sources of variation (see Table 1).

Traditionally, (behavioral) engagement has been considered a between-*student* variable. For example, Jack might typically be more engaged in class than Sophie – no matter the teacher or situation. The majority of studies presented in Table 1 seem to support this assumption. Variance decompositions of (behavioral) engagement find very high proportions of between-student variance (81–93%) if the student level is not further disaggregated. However, if within-student variability is modeled, between-student variance estimates are lower – though still considerable (15–85%; Martin et al., 2015, 2019; Motti-Stefanidi et al., 2015; Patall et al., 2016; Pöysä et al., 2018; Ruzek & Schenke, 2019; Shernoff et al., 2016; Vasalampi et al., 2016). Designs that do not further disaggregate the student level (Jang et al., 2010; Marks, 2000; Nie & Lau, 2009; Reyes et al., 2012) therefore potentially overestimate the role of stable between-student factors in (behavioral) engagement: behavioral engagement and engagement in general thus may be less trait-like and more malleable than has been assumed (Ruzek & Schenke, 2019).

In our study, therefore, we have included a within-student measurement of behavioral engagement. Especially given the growing conceptual emphasis on the situational nature of behavioral engagement, inclusion of intra-individual variation or a *situational* level seems essential for a more accurate estimation of stable differences between students (Gilbert, Petscher, Compton, & Schatschneider, 2016). Jack might be more engaged after an interaction with his math teacher who tells him what to do when trying to solve a particular math problem than after an interaction with another teacher in a different class, working on a different task at a different time of the day. As shown in Table 1, behavioral engagement appears to vary quite extensively from situation to situation, especially within days (variance estimates ranging between 12 and 74%; Martin et al., 2015, 2019; Pöysä et al., 2018; Shernoff et al., 2016; Vasalampi et al., 2016). When studies do not model this part explicitly, situational variability is typically confounded with (i.e., 'hidden' in) the student level (Mainhard et al., 2018). Also, none of the studies in Table 1 measured students' engagement from moment-to-moment, although moment-to-moment changes within lessons potentially induce corresponding moment-to-moment fluctuations in engagement. At the teacher level (or the teacher-class level), there is evidence for stable but relatively small differences between teachers: estimates vary around 10% (see Table 1).

Previous studies did not study whether there is behavioral engagement variation that could be ascribed to the *teacher-student dyad*. Behavioral engagement has been shown to be prone to dyadic influences such as the quality of teacher-student relationships (Birch & Ladd, 1997) and the nature of teacher-student interactions (Cadima, Doumen, Verschueren, & Buyse, 2015). Empirical evidence of relationship effects on (behavioral) engagement seems to be nonexistent. For a related concept, student enjoyment, evidence shows that students' positive and negative emotions in class "are not only associated with teacher trait-like characteristics and students' dispositions (i.e., stable differences between teachers and between students) but also depend on how teachers adjust their teaching to specific students and classes" (Mainhard et al., 2018, p. 117). Thus, engagement levels can be expected to not only depend on student's tendencies to be engaged to a certain extent or teachers' tendencies to induce a certain level of engagement, but also on the specific match between a particular student and a particular teacher (i.e., a relationship or dyadic effect).

Based on the above, in our study we decided to investigate the extent to which *situation, student, teacher,* and *relationship* effects contribute to variation in behavioral engagement in the context of instructional support during seatwork. Modelling these levels together provides the unique opportunity to disentangle behavioral engagement variation from moment to moment. Yet, such descriptive analysis falls short of indicating avenues for interventions to optimize engagement in school (Pöysä et al., 2018). If variation in behavioral engagement is indeed attributable to situational and relational (i.e., dyadic) factors, we expect that variability in engagement can partly be explained by teaching practices and specifically instructional teacher-student interactions (Pöysä et al., 2018).

4. Adaptive support and behavioral engagement

The effectiveness of teacher-student instructional interactions during seatwork hinges partly on the adaptiveness of the instruction or support that it provides. Adaptive support (van de Pol, Mercer, & Volman, 2019; van de Pol, Volman, Oort, & Beishuizen, 2015; Wood, Wood, & Middleton, 1978) ensures that students know what to do and how to proceed independently (van de Pol et al., 2015). One example of adaptive support would be when a teacher asks an open question that fits a student's current understanding and at the same time stimulates that

Table 1
Overview of variance decompositions of (behavioral) engagement in previous studies.

			Variance at different levels of analysis						
			situational (variation in time)						
Engagement type	Article	Construct measurement	within-day	day	week	months/ years	student	class/ teacher	school
Behavioral engagement	Jang, Reeve, & Deci (2010)	6 observer-rated engagement-related student behaviors at the classroom level: attention, effort, verbal participation, persistence, positive emotion, and voice	n.a.	n.a.	n.a.	n.a.	93%	7%	n.a.
	(Motti-Stefanidi, Masten, & Asendorpf (2015)[1]	absenteeism (A; obtained from school record) and 6 teacher-rated items assessing degree of student motivation and engagement in school work (teacher-rated engagement; E)	n.a.	n.a.	n.a.	73% (A) 44% (E)	19% (A) 45% (E)	8% (A) 11% (E)	n.a.
	(Patall, Vasquez, Steingut, Trimble, & Pituch (2016)	3 student-rated items: working hard (W), participation (P), paying attention (A)	n.a.	59% (W) 52% (P) 57% (A)	n.a.	n.a.	37% (W) 43% (P) 38% (A)	4% (W) 4% (P) 5% (A)	n.a.
	Ruzek & Schenke (2019)	5 self-reported items assessing student's participatory behaviors in class (based on Wellborn's (1991) behavioral engagement and disaffection scales)	n.a.	n.a.	n.a.	85%	15%	n.a.[2]	n.a.
Engagement in general	Jang, Reeve, & Deci (2010)	4 self-reported individual student items: paying attention (behavioral), working hard (behavioral), trying to learn as much as possible (cognitive), and enjoyment (emotional)	n.a.	n.a.	n.a.	n.a.	86%	14%	<1%
	Marks (2000)	4 student-rated component measures (one factor): student effort, attentiveness, lack of boredom in class, completing class assignments; measured at elementary (E), middle (M), and high (H) school levels	n.a.	n.a.	n.a.	n.a.	84% (E) 88% (M) 92% (H)	12% (E) 9% (M) 8% (H)	4% (E) 3% (M) - (H)
	Martin, Papworth, Ginns, Malmberg, Collie, & Calvo (2015)	3 student-rated measures of (mal)adaptive engagement: persistence (P), self-handicapping (S), disengagement (D)	27% (P) 12% (S) 35% (D)	2% (P) 1% (S) 1% (D)	6% (P) 4% (S) 8% (D)	n.a.	65% (P) 83% (S) 56% (D)	n.a.	n.a.
	Martin, Mansour, & Malmberg (2019)	student-rated Motivation and Engagement Scale – High School (MES-HS) (Martin, 2011): adaptive engagement (planning behavior, task management, persistence) and maladaptive engagement (self-handicapping, disengagement)	33%[3]	2%	3%	n.a.	62%	n.a.	n.a.
	Nie & Lau (2009)	students' report of their attention, effort, and participation in classroom activities (one factor)	n.a.	n.a.	n.a.	n.a.	92%	8%	n.a.
	Pöysä, Vasalampi, Muotka, Lerkkanen, Poikkeus, & Nurmi (2018)	3 latent factors of engagement: behavioral/cognitive engagement (B), emotional engagement (E), disaffection (D)	63% (B) 70% (E) 53% (D)	7% (B)[4] 8% (E) 8% (D)	n.a.	n.a.	29% (B) 22% (E) 39% (D)	n.a.	n.a.
	Reyes, Brackett, Rivers, White, & Salovey (2012)	adaptation of Engagement vs. Disaffection Scale (Furrer & Skinner, 2003): students' self-reported perceptions of effort, interest, and enjoyment while initiating/sustaining learning activities	n.a.	n.a.	n.a.	n.a.	81%	19%	n.a.
	Shernoff et al., 2016	composite of three student-rated items: enjoyment, concentration, interest	15%[5]	n.a.	n.a.	n.a.	40%	12.6%	n.a.
	Vasalampi, Muotka, Pöysä, Lerkkanen, Poikkeus, & Nurmi (2016)	five-factor structure of engagement: behavioral engagement, emotional engagement, competence experiences, disaffection, help seeking (measured by 17 items)	54–74%[6]	n.a.	n.a.	n.a.	26–46%	n.a.	n.a.

[1] This longitudinal study reports *development* over years, which distinguishes it from the other, non-longitudinal studies.
[2] In a footnote, the authors report that the person-level ICCs did not change in a three-level unconditional variance components model including an additional classroom level (ICC of behavioral engagement at the classroom level was 0.06).
[3] The displayed figures are average percentages of all factors (see description).
[4] Due to the "small ICCs found between days of week and the study design which involved only one week of measurement for each participant", Pöysa et al. decided to construct a two-level hierarchical model with days included as dummy coded variables instead of a three-level model with days as an independent variable (2017, p. 68).
[5] "Instructional episode", i.e. (part of) a lesson. Shernoff et al. (2016) have used two-level HLM regression models, but also report about a teacher (12.6%) and observation (% not mentioned) level.
[6] ICCs are reported per item, which accounts for the variance range.

understanding towards a next step. Another example would be when a teacher provides an explanation if the student cannot solve a problem on their own (van de Pol & Elbers, 2013). If support is non-adaptive, on the other hand, students are either underchallenged (i.e., the level of support lies below the students' level of understanding) or overchallenged (i.e., the level of support lies beyond the student's level of understanding). An example of that would be when a teacher provides an answer or explanation when the student already understands the topic or when a teacher asks an open question that is far beyond a student's current understanding (van de Pol & Elbers, 2013). In both cases, students are unlikely to stay on task because they may either get bored or feel unable to finish the task. Therefore, support that adaptively zooms in on the student's (mis)understanding is crucial in this context.

Previous research provides initial evidence that adaptive support relates to behavioral engagement. Based on students' responses to questionnaires on perceived teaching quality, which included measures of adaptive support, and self-assessed behavioral engagement, operationalized as effort regulation, Leon, Medina-Garrido, and Núñez (2017) observed that teaching quality predicted the level of displayed behavioral engagement at the class level. Also, in a hierarchical regression model of students' time-on-task during small-group work, Chiu (2004) found that students were more likely to engage with the task after more adaptive teacher interventions. Similarly, Lutz, Guthrie, and Davis (2006) observed during whole-class teaching that at class level, whole-class directed adaptive support and individualized teacher prompts were effective to sustain engagement in complex tasks. Based on these findings, we expect that teacher adaptive support has the potential to predict student engagement at several levels (e.g., student, class, teacher) and specifically at the situational level. Although previous studies did not concern seatwork contexts, we expect that adaptive support is positively associated with engagement in the seatwork context. Other than the studies just described, we operationalize adaptive support at the situational level at which we find the highest variance in engagement, which potentially allows us to explain the largest share of the attested variance.

5. The present study

This study aims to address the situational nature of behavioral engagement in the context of independent seatwork. We hope to extend current understanding of behavioral engagement by taking an observational approach to situational, student, teacher, and student–teacher components of behavioral engagement displayed by students around teacher-student interaction in seat-work settings in different classrooms and at multiple time points. First, we decomposed variability in student engagement. We viewed situational engagement as being nested within two higher levels, the student and teacher in question (i.e., cross-classification), which allowed us to estimate the degree to which dyadic or relationship effects play a role in student engagement. Based on evidence of the importance of situational effects reported in Table 1 (Martin et al., 2015; Patall et al., 2016; Pöysä et al., 2018; Shernoff et al., 2016; Vasalampi et al., 2016) and indications of the importance of relationship effects to student-related variables (Mainhard et al., 2018), we expected that inclusion of a situational and relationship level would reduce between-student variability in comparison to earlier studies of student engagement. In line with that, we expected that observational measures are less person-dependent compared to student self-report measures, which possibly also reduces student-related variation in engagement levels.

Second, based on the variance composition in engagement, we estimated the extent to which variation in behavioral engagement can be explained by teachers' adaptive support. Given previous evidence for a link between adaptive support and behavioral engagement (Chiu, 2004; Leon, Medina-Garrido, & Núñez, 2017; Lutz et al., 2006), we hypothesized that variation in adaptive support would relate positively to variation in behavioral engagement during independent seatwork.

6. Method

6.1. Data and participants

We analyzed video recordings of seatwork situations, visiting 150 lessons of four prevocational education classes at four different Dutch schools (virtually all lessons students attended during the two weeks). At this grade-level, almost all subjects are taught by a different teacher. Two video cameras were installed in the classrooms and teachers were asked to wear microphones and students were recorded with audio recorders put at their desks (one recorder per two or four students). Teachers were asked to organize their lesson such that independent seat work was a major activity but did not receive any specific instructions or training in providing adaptive support.

Students in our dataset (56 male, 31 female) were in four different classes, all in their first year of secondary education (aged 11, 12, or 13). Teaching experience of the participating teachers ($N = 37$) varied between 1 and 35 years ($M = 12.2$, $SD = 10.4$).

6.2. Procedure

Teacher-student interactions were included for analysis only if: the interaction was task-related ($N = 463$), the interacting dyad interacted more than once to allow for partitioning of situational variance (i.e., variation in time; Kenny et al., 2001; $N = 330$), and at least one of the dyad members involved in the interaction participated also in another dyad, to allow for variance estimation at the student and teacher level. The final sample included 324 interactions between 107 unique student–teacher dyads ($M = 3.0$ per dyad, $SD = 1.4$, min. = 2, max. = 10). Selected interactions occurred during mathematics classes ($N = 112$), English language classes ($N = 72$), Dutch language classes ($N = 45$), or one of the other classes ($N = 95$). The average duration of interactions was 54.8 s (min. = 3.8, max. = 371.2, $SD = 47.9$).[2]

6.3. Measures

Behavioral Engagement. Episodes of independent seatwork preceding and following the interaction were coded for behavioral engagement. In line with Lee and Anderson (1993), Lutz et al. (2006), and Skinner, Kindermann, and Furrer (2009), we operationalized behavioral engagement as on-task behavior. We observed on-task behavior from second to second instead of obtaining student- or teacher-ratings to capture even minor moment-to-moment fluctuations that cannot be captured in interval-measures of on-task behavior generally used in educational research (Lutz et al., 2006). A comprehensive description of our coding approach is attached as Appendix A.

Student's on-task behavior was coded per second in Mediacoder 2009. Both before and after the interaction, the duration of the on-task behavior coding was bounded by other teacher-student interactions (involving the target student) or teacher's whole-class comments. If such events did not occur, fragments were coded to a maximum of two minutes preceding or following the focus interaction. Only fragments of at least 1 s were included, resulting in 252 pre-interaction ($M_{\text{duration}} = 85$

[2] The selection of instructional support interactions resulted in a select sample of students included in the analyses. To check whether the students included in the current sample were different from those not included (i.e., because they did not have content-related interactions with their teachers in the observation period), we tested whether the groups differed on a student-rated survey on general behavioral engagement in class (Skinner et al., 2009; 10 items, 4-point scale) and on four achievement-related measures (final grade for the core subjects mathematics, English language, and Dutch language, and their scores on a standardized school leavers test administered at the end of primary school). Independent samples t tests showed no significant differences between the groups on any of the measures.

sec., $SD_{duration}$ = 42 sec., median = 112 sec., range 3–120 sec., 120 fragments of maximum length of 120 sec.) and 252 post-interaction ($M_{duration}$ = 81 sec., $SD_{duration}$ = 45 sec., median = 113 sec., range 1–120 sec., 117 fragments of maximum length of 120 sec.) on-task fragments.

On-task behavior was coded with four codes based on behavioral features mentioned in engagement studies (see Appendix A, Table A1): on-task behavior, off-task behavior (e.g., non-content related interaction with fellow students), creating conditions for on-task behavior (e.g., waiting to ask a question with hand raised), and uncodable (e.g., student off-screen). A student was considered on-task when (s)he was attentive to and concentrated on (Skinner et al., 2009) his/her own task, either actively or passively (e.g., working or listening; Volpe, DiPerna, Hintze, & Shapiro, 2005), or when (s)he was involved in activities expected or instructed by the teacher (Imeraj et al., 2013).

Coding was conducted by the first author and a research assistant. To ensure reliable coding, coders first pilot coded several randomly chosen on-task fragments; anything unclear was settled before coding began. Next, 10% of the pre- and post-interaction fragments (randomly selected) were double coded. Second-by-second agreement (Bakeman & Quera, 2011) was calculated in two steps. First, we determined whether coders agreed on which parts of the fragments had to be coded following the rules displayed in Fig. A1 in Appendix A. All content-related teacher-student interactions were coded, regardless of the position of the teacher relative to the student. These included configurations where a teacher was standing next to or approaching the student's table, as well as situations where the teacher was located at the front of the classroom talking to a student in the back of the class. For the code/do not code decision, Gwet's AC1 (an agreement measure for nominal scales recommended over alternative coefficients (Gwet, 2014) was 0.86. Second, we determined whether coders agreed on codings of to be coded fragments. Gwet's AC2 (the weighted AC1 to be used for non-nominal data) for these codings was 0.91. Coding disagreements were discussed and resolved by the coders after interrater reliability calculation, resulting in an additional nine coding rules (see Appendix A). Since interrater reliability for both steps is considered very good (Landis & Koch, 1977), subsequent coding was continued separately.

For each pre- and each post-interaction on-task fragment, an interval score of on-task behavior was calculated by dividing the duration of the parts of the fragment coded as on-task by the total duration of the coded fragment (e.g., 80 s on-task of a total of 120 s would yield a score of 0.67).

Adaptive Support. The 324 selected teacher-student interactions were coded for adaptive support using an existing coding scheme (van de Pol et al., 2019), see also (van Halem, 2016). Teacher-student–teacher sequences (i.e., three-turn sequences) were coded in three rounds (see also Appendix B). First, teacher turns were coded for diagnostic strategies: a teacher turn was considered a diagnostic strategy if it functioned to obtain information about student understanding. These turns were excluded from further analyses (except if accompanied by support in the same turn), because we focused on adaptive *support* and diagnostic strategies are not considered support (e.g., (van de Pol, Volman, Oort, & Beishuizen, 2014). Second, all teacher turns were coded with regard to the degree of control or regulation the teacher exercised in his or her support. This control was coded on two dimensions: 1) with regard to the form of the interaction (no question was considered to be the highest degree of control, a closed question medium degree of control, and an open question the lowest degree of control), and 2) with regard to the content of the interaction (highest control level: support provides direct new content, i.e., instructions or explanations, or provides the opportunity for imitation; medium control level: support contains partial new content, through explicit clues, suggestions, tips, or feedback; or low control level: support contains no new content; see Appendix B). Third, student turns were coded for accuracy of student understanding; coding student understanding is necessary to establish the actual degree of adaptivity of teachers' support.

Interrater reliability was established using 52 teacher turns and 52 student turns selected from 25 interactions (for selection process details, see (van Halem, 2016). Reliability for the coding rounds was substantial to almost perfect: $AC1_{\text{Diagnostic Strategy}} = 0.93$, $AC2_{\text{control-form}} = 0.92$, $AC2_{\text{Control-content}} = 0.73$, $AC2_{\text{Accuracy of Student Understanding}} = 0.87$ (Gwet, 2014; Landis & Koch, 1977). Occasional differences between codings were examined closely and codings were adapted accordingly to further ensure reliability of future coding.

To determine whether a three-turn (teacher-student-teacher) sequence was adaptive, two final adaptivity scores were assigned to each sequence based on the codes assigned to the first teacher turn, student turn, and second teacher turn; one adaptivity score (0 or 1) represents the degree to which the sequence was adaptive with regard to its form, the other adaptivity score (0 or 1) represents the degree to which the sequence was adaptive with regard to the content. Generally, increasing control (form and content) upon low student understanding and decreasing control (form or content) upon high student understanding was considered adaptive (Wood et al., 1978). One sequence could thus receive an adaptivity score of 0 (non-adaptive on both dimensions), 1 (adaptive on one dimension), or 2 (adaptive on both dimensions), see van de Pol et al. (2019). These sequence adaptivity scores were aggregated to the interaction level.

An example coded interaction is presented in Appendix C.

7. Analysis

The first analytic step, variance decomposition, was conducted for pre- and post-interaction engagement separately. Since our aim was to distinguish between stable teacher, student, dyadic and situational effects, we employed a cross-classified multilevel model. Pre- and post-interaction engagement were modelled as nested within two levels, namely within the teacher and the student, and the dyadic teacher-student effect was modeled as the interaction term between these two levels (Fielding & Goldstein, 2006). In line with Kenny, Kashy, and Cook (2006) the student and teacher level can be understood as reflecting "the consistency of an actor's behavior across interaction partners" (p. 129). The interaction between the teacher and student level is reminiscent of what Kenny has referred to as the relationship or dyadic effect (i.e., reflecting the amount of variance in student engagement due to specific adjustment in student–teacher pairings). Finally, estimation of the situational effect allows us to segregate the extent to which behavioral engagement is consistent between two interactions of a specific student–teacher dyad from any instabilities between the interactions due to variation in time or chance fluctuations. Controlling for mean-level differences by adding school/class as dummies did not make any difference to the models' estimates.

Second, to investigate the predictive value of the quality of teacher-student interactions in terms of teacher adaptive support, we first aggregated adaptive support to the levels at which variance in post-interaction engagement was found. The reliability of these aggregates was evaluated by calculating ICC1 and ICC2 (Snijders & Bosker, 2011). If reliable, the aggregates were added to the model at the appropriate level.

All analyses were performed using SPSS 25. Syntax (Albright & Marinova, 2010; Kenny, 2007) of each analysis is presented in the online supplemental material.

8. Results

8.1. Preliminary analyses

Residuals at all levels of the variance decompositions were approximately normally distributed, showing only some slight deviations at the extremes. Behavioral engagement analyses were based on 252 pre-interaction on-task fragments and 252 post-interaction on-task fragments. On average, 50% of the duration of a pre-interaction on-task fragment ($M = 0.498$, $SE = 0.050$) and 71% of the duration of an on-task fragment ($M = 0.707$, $SE = 0.040$) was coded as displaying on-task behavior. A paired sample t-test showed that pre-interaction behavioral engagement (M = 0.504, SD = 0.362, min. = 0, max. = 1) was considerably lower than post-interaction behavioral engagement (M = 0.693, SD = 0.306, min. = 0, max. = 1), $t(225) = 7.42$, $p < .001$). On average, teachers showed moderate adaptive support ($M = 0.99$, $SD = 0.41$, min. = 0.00, max. = 2.00).

8.2. Variance decomposition of behavioral engagement

Variance component estimates of pre-interaction behavioral engagement at the student, teacher, dyadic, and situational levels are presented in Table 2. Variance in pre-interactional engagement was situational (situation level and error, 73%, $p < .001$) and teacher-related (27%, $p = .018$). No student-related variation was attested.

Variance component estimates of post-interaction behavioral engagement at the student, teacher, dyadic, and situational levels are presented in the first model in Table 3.

Variance in post-interaction behavioral engagement was mainly due to variation in time (situation level and error, 68%, $p < .001$). Teacher-level variance, though non-significant (17%, $p = .083$), was slightly larger than student-level variance in post-interaction behavioral engagement (15%, $p = .043$). No variance was situated at the dyadic level[3].

To control for initial levels of behavioral engagement, we first aggregated pre-interaction behavioral engagement to the levels at which variance in post-interaction engagement was found (student, teacher, situation) and evaluated the reliability of these aggregates using ICC1 and ICC2 (Snijders & Bosker, 2011 see Table 4).

The aggregate at the teacher level was reliable, so pre-interaction behavioral engagement was entered at both the situational/error level and the teacher level (see M2 in Table 3). Pre- and post-interaction behavioral engagement residuals were linearly related. As follows from Table 3, pre-interaction behavioral engagement explains part of the variability in post-interaction behavioral engagement at the teacher-level, in addition to a smaller portion of situational and error variance. The model controlling for pre-interaction behavioral engagement fitted significantly better than the intercept-only model ($\chi^2 = 84.651 – 50.449 = 34.202$, df = 1, $p < .001$).

Table 2
Parameter estimates (and standard errors) for pre-interaction behavioral engagement.

Parameter	B (SE)
Fixed effects	
Intercept	0.498 (0.050)*
Random parameters	
Student variance	0.000 (0.000)
Teacher variance	0.036 (0.015)*
Dyadic variance	0.000 (0.000)
Situational + error variance	0.095 (0.009)*
−2LogLikelihood	156.023
AIC	164.023
BIC	178.125

* Significant at α = 0.05. Significance is based on Wald-Z tests.

[3] Given the absence of any dyadic variance, we tested whether the dyadic model could be removed from the model. Removal of the dyadic level from M1 resulted in only minor and non-significant model fit changes. Therefore, and given the conceptual importance of the level in the context of the current study's focus on, amongst others, the dyadic effect, we decided to leave the dyadic level in the model.

Table 3
Parameter estimates and standard errors for models of post-interaction behavioral engagement.

Parameter	B (SE)			
	M1: Intercept-only	M2: Covariate	M3: Adaptive support	M4: Duration of fragments
Fixed effects				
Intercept	0.707 (0.040)*	0.425 (0.073)*	0.409 (0.086)*	0.466 (0.093)*
Pre-interaction behavioral engagement				
at teacher level		0.444 (0.143)*	0.440 (0.149)*	0.451 (0.148)*
at situation + error level		0.115 (0.056)	0.110 (0.058)	0.105 (0.058)
Adaptive support			0.023 (0.045)	0.018 (0.045)
Duration of post-interaction fragment				−0.001 (0.000)
Random parameters				
Student variance	0.014 (0.007)*	0.015 (0.007)*	0.016 (0.007)*	0.016 (0.007)*
Teacher variance	0.016 (0.009)	0.005 (0.004)	0.006 (0.004)	0.005 (0.004)
Dyadic variance	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Situational + error variance	0.065 (0.007)*	0.058 (0.006)*	0.056 (0.006)*	0.055 (0.006)*
−2LogLikelihood	84.651	50.449	46.991	58.402
AIC	92.651	58.449	54.991	66.402
BIC	106.753	72.077	68.244	79.635

* Significant at α = 0.05. Significance is based on Wald-Z tests.

Table 4
Intraclass correlations for pre-interaction behavioral engagement.

Level	ICC1 (p)	ICC2[1]
student	0.00 (–)	n.a.
teacher	0.27 (0.018)	0.86
dyad	0.00 (–)	n.a.
situation + error	0.73 (<0.001)	n.a.

[1] Only ICC2s of aggregates at levels at which post-interaction behavioral engagement variation was found were calculated.

9. Prediction of behavioral engagement with adaptive support

To establish how much of the behavioral engagement variance at the different levels can be explained by variation in adaptive support, we first aggregated the adaptive support measure to the levels at which post-interaction behavioral engagement variation was significant, i.e., the situational level, the student level, and the teacher level. Then, we established whether adaptive support could be reliably assessed at those levels, by calculating the ICC1 and ICC2 (Snijders & Bosker, 2011); see Table 5). As there was no variance in adaptive support at the student level, adaptive support could not be reliably aggregated to that level.

ICC2 for the teacher-level aggregate was calculated using the mean

Table 5
Intraclass correlations for adaptive support.

Level	ICC1 (p)	ICC2[1]
student	0.00 (–)	n.a.
teacher	0.05 (0.330)	0.45
dyad	0.08 (0.172)	n.a.
situation + error	0.87 (<0.001)	n.a.

[1] Only ICC2s of aggregates at levels at which post-interaction behavioral engagement variation was found were calculated.

number of interactions per teacher, which is the usual approach if the number of observations per unit varies (Lüdtke, Robitzsch, Trautwein, & Kunter, 2009). Based on the mean number of interactions per teacher, the teacher-level aggregate was not reliable (0.045). Given these results and the large amount of variability in student engagement at the situational level, adaptive support was entered into model M2 as a predictor only at the lowest level.

Post-interaction behavioral engagement and adaptive support residuals were linearly related. Parameter estimates are presented in model M3 in Table 3. Adaptive support did not significantly explain behavioral engagement variance in model M3 ($\chi^2 = 50.449–46.991 = 3.458$, df = 1, $p = .063$). After controlling for pre-interaction behavioral engagement, adaptive support only explained a mere 2% of the situational and error variance.

Given the large variety in duration of post-interaction fragments, we exploratively investigated the relation between degree of on-task behavior after the teacher-student interaction and the length of the coded fragment. The duration of post-interaction fragments was indeed correlated to the degree of on-task behavior during that fragment ($\rho = -0.018$, $p = .004$, $N = 252$). Model fit, however, decreased significantly after addition of the duration of post-interaction fragments as a predictor to model 3 ($\chi2 = 46.991–58.402 = -11.411$, df = 1, $p < .001$), which is why we chose the third model as the final model.

10. Discussion

Engagement in school is a complex, but particularly relevant construct in the field of educational psychology (Sinatra, Heddy, & Lombardi, 2015). We explored its complexity by studying variability in behavioral engagement in a common school setting, independent seatwork. We investigated to what degree behavioral engagement in the context of instructional support during seatwork should be conceived as a situational, dyadic, teacher or student construct and studied its relation with provision of adaptive support by teachers. The analyses point at two main findings: engagement varied mainly from moment to moment while there was no evidence for dyadic teacher-student effects in student engagement, reflecting their interpersonal adjustment, and secondly, the degree to which teacher support in dyadic teacher-student interactions was adaptive did not predict student engagement levels right after these interactions.

First, *engagement in the context of instructional support during seatwork (both pre- and post-interaction) varied mainly from moment to moment.* Despite our selection of dyadic situations (i.e., teacher-student interaction), no variation in engagement on the dyadic level was found. Being situational, the attested variation might be considered as variation in the complex combination of divergent contextual aspects over time (e.g., co-seated students, see Santoyo, Jonsson, Anguera Argilaga, and López Lopez (2017), type of task) (Kenny, Mohr, & Levesque, 2001). Although the lowest-level includes residual variance, we expect error variance not to be very extensive given the high reliability of coding and given the particularly malleable nature of behavioral engagement (Fredricks & McColskey, 2012). The attested situational variance points at the local nature of engagement: even though behavioral engagement was measured in a specific situation (i.e., during seatwork after a teacher-student interaction about the task the student was working on), students' engagement varied from one situation to the next (Fredricks & McColskey, 2012).

The assumption that (behavioral) engagement is a student characteristic has been based on findings of previous variance decomposition studies of (behavioral) engagement. These studies however generally took the student as their lowest investigated level (Jang et al., 2010; Marks, 2000; Nie & Lau, 2009; Reyes et al., 2012), which are typically confounded with situational variance. Our current findings suggest that student-level variance partially shifts 'down' if lower levels are included (cf. Martin et al., 2015; Patall et al., 2016). Apt modeling and choice of measures is thus crucial. If behavioral engagement in the context of instructional support during seatwork would indeed be largely context-dependent and would arise locally in interaction, micro-level measures of engagement in interaction aid us in discovering how situational factors contribute to fluctuations in engagement from moment to moment. General measures of overall engagement in school, for example based on surveys, would not suit that aim. Rather, those general measures of school engagement are the method of choice when one aims to predict student outcomes such as school drop-out and academic achievement (Fredricks et al., 2004); as such, they could point at general engagement patterns that may help us to understand why a certain student is engaged in a certain context at a certain point in time. Both types of measures have their own validity issues: measures based on student- or teacher-rated engagement are subject to potential subjective biases (Briesch, Chafouleas, & Riley-Tillman, 2010). Measures based on observations may not be able to differentiate between students who are highly engaged, and students who appear to work on their task but think about something else—or vice versa (Fredricks & McColskey, 2012; Peterson, Swing, Stark, & Waas, 1984). Given the observable nature of behavioral (as compared to cognitive and emotional) engagement (Suárez-Orozco, Suárez-Orozco, & Todorova, 2009) and the potential impact rater bias in self-report measures would have had on the results of our variance decomposition study (Briesch et al., 2010), we believe that our observational approach has helped to gain more insight in students' on task behavior. As students' behavioral engagement appeared to be fairly independent of the student themselves and their teacher, such investigation potentially yields widely applicable advice for raising engagement in that classroom situation.

While evidence for the situational nature of engagement in seatwork setting was abundant, we found no evidence for dyadic variation in behavioral engagement in this setting. Although behavioral engagement has been shown to be prone to dyadic influences (Birch & Ladd, 1997; Cadima et al., 2015) and is generally described as a dynamic construct open to influences from the teacher-student relationship (Lutz et al., 2006), the current results suggest that engagement behavior in itself might not be dyadic in nature. This distinguishes engagement behavior from, for example, interpersonal perception, which inherently includes two people (Kenny et al., 2006). In Pianta et al.'s (2012) understanding, engagement is *contingent on* longer-term, relational processes and "their behavioral expression in interpersonal interactions in the classroom" (p. 366). In that sense, non-interactional features like long-term relationships could be seen as *influencing interaction* (see also Martin & Collie, 2019; Vollet, Kindermann, & Skinner, 2017). Yet, our data suggest that, even if measured in dyadic situations such as directly following teacher-student interaction, variation in engagement is independent of the specific teacher-student dyad involved in that interaction. Whether this finding holds in other contexts where teacher-student interaction is more frequent or long-term is a topic for further exploration. Such exploration could model observations with self-report measures in a comparative analysis to discern to what extent the lack of relationship effects on engagement is an artifact of our observational approach (Azevedo, 2015).

It is important to note that, from our observational approach, variation in pre-interaction engagement was quite strongly teacher-related, in contrast to between-teacher effects reported earlier (e.g., Jang et al., 2010; Patall et al., 2016). This difference could be the result of the context of our observations: instructional support. These situations center around a cognitive and interactional exchange between teacher and student, focusing less on teacher characteristics like interpersonal behavior. Also, the difference with previous literature could be the result of observations of differences between teachers in the way they organize seatwork. In some classes, students were expected to raise their hand to summons their teacher and wait for them to come to their table. We coded this behavior as creating conditions for on-task behavior, not on-task behavior per se. If students would have been able to continue working while otherwise signaling they needed help on some aspect of their work, that could have resulted in more behavior coded as on-task. More generally, the difference between pre-interaction and post-interaction variability shows that variance decompositions of behavioral engagement can be dependent on the specific situation observed.

Based on the variance decomposition findings, we could determine the appropriate level for adaptive support as a predictor of variation in engagement. Our second main finding, then, is that *situational variation in behavioral engagement could not be explained by adaptive support* – despite the slight increase in model fit on addition of the adaptive support predictor. Contradicting previously mentioned indications of a positive relation between adaptive support and engagement (Chiu, 2004; Leon et al., 2017; Lutz et al., 2006), the current data provide no evidence that the degree of adaptivity of support in teacher-student interaction during seatwork is related to differences in observed behavioral engagement after that interaction. Possibly, it is not so much the *adaptivity* of the support that contributes to behavioral engagement. After all, support established between teacher and student is simply one aspect of the complex instructional context that gives rise to a particular level of engagement at that moment for that student (Taboada Barber, Buehl, & Beck, 2017). Recent evidence points at the importance of, for example, quality feedback, which, in conjunction with other teacher practices was found to predict behavioral engagement in the classroom (McKellar, Cortina, & Ryan, 2019). An alternative explanation for the lack of evidence for a link between behavioral engagement and adaptive support may be found in the lack of variation between teachers' adaptive support in this sample. Also, a methodological issue might be at play here. While the duration of our fragments following teacher-student interaction might have been long enough for interactional processes to affect post-interaction behavior and short enough to not include all types of post-interaction classroom incidents not related to the interaction, it is possible that effects of adaptive support show later than two minutes post-interaction. Close sequential examinations of teacher and student behavior during and following teacher-student interaction would be needed to shed light on this question.

Although we found no evidence for a relation between adaptive support and behavioral engagement, teacher-student interaction seems to be helpful in some way: absolute levels of engagement behavior increased substantially from pre- to post-interaction. This supports the idea that teacher-student interaction indeed aids students in proceeding with the task at hand (i.e., to be engaged; Chiu, 2004; van de Pol et al., 2015). Such aid could hinge on the content of the interaction, which is supposed to help students do what they are expected to do. Alternatively, the interpersonal contact during the interaction could motivate and engage students to get back to work, persist in trying to solve hard tasks, or otherwise meet what is expected from them. It is noteworthy that most interactions between students and their teachers in our larger dataset were not related to the content (and therefore could not be coded for adaptivity). Thus, during seatwork, teachers already do a lot of process-related work, perhaps aiming to increase on-task behavior.

Reflecting on our findings, several analytic issues warrant careful interpretation of the results. First, differences between the current findings and earlier findings (e.g., those reporting considerable student-related variation; Jang et al., 2010; Marks, 2000; Martin et al., 2015, 2019; Nie & Lau, 2009; Patall et al., 2016; Reyes et al., 2012; Shernoff et al., 2016) may be an artifact of differences in methodological approaches. The current observational approach taps into moment-to-moment engagement *as it happens*, viewed from an observational perspective – while commonly applied self-report approaches tap into engagement levels as perceived by the students themselves. Observational measures circumvent commonly acknowledged bias issues that affect student's self-reports (Azevedo, 2015; Fredricks, Filsecker, & Lawson, 2016), but they may also over- or underestimate the quality of engagement in terms of, for example, effort and participation (Fredricks & McColskey, 2012). The differences between findings from self-report studies and observational studies may thus be attributable to the degree to which these approaches capture engagement processes in real time. Alternatively, the specifics of the seatwork context may bring to the fore features of engagement that may be hidden or not as pronounced in other contexts. Engagement during independent seatwork might be relatively susceptible to task characteristics, such as the clarity of the task, and teacher-related aspects, such as clarification of the task and good classroom management (see Anderson, 1984; Helmke & Schrader, 1988). These features are less pronounced in, for example, whole class instruction settings. Also, the granularity of our observational measure (Azevedo, 2015) is much smaller than the observational measurements reported in Jang et al. (2010), who found large amounts of student-related variability in engagement. The current challenge in engagement research is to integrate different approaches to foster conceptual as well as practical understanding of the complexities of the construct (Azevedo, 2015; Fredricks et al., 2016).

Second, in our analysis, we coded a request for support (for example, by a raised hand) as *creating conditions* for engaged behavior, not as on-task behavior per se. Arguably, asking for help, as a way to find information that allows the student to proceed with the task, is already indicative of engagement with the task. At the same time, however, having one's hand raised while talking to a fellow student about non-task-related topics clearly displays disengagement with the task – making coding of such behavior as engaged or disengaged rather ambiguous at times. Possibly, coding degrees in engagement rather than dichotomous labeling of the behavior as either engaged or not engaged could address this issue. Such coding could also allow for differentiation between different types of engaged behavior, which might foster understanding of how the quality of engagement varies over time (and, possibly, between individuals).

Third, a substantive part of all students observed was not included in the data for analysis. These students did not have content-related interactions with their teachers during the observed classes. Although their levels of engagement could be informative to understanding engagement variation in general, measures of their engagement would be less relevant to understanding the factors that are at play in the particular *interactional* setting that we chose to focus on: engagement in the context of instructional support during seatwork. Still, understanding the variation in their engagement, too, would be an issue of interest for further study. If these students did not need support, did need but did not ask for support, or were not approached with support, how did that affect their engagement? Also, they did not interact with their teachers about the content of the lesson – but did they have other types of interactions, and how is this related to their engagement levels? These questions are outside the current study's scope, but their answers could definitely help us find another piece of the dynamics of behavioral engagement puzzle.

In conclusion, the current study shows that behavioral engagement in an independent seatwork context likely is more than the sum of stable teacher and student characteristics alone. The analyses contribute to our understanding of engagement in several ways. Theoretically, our findings underscore the situational nature of engagement. The data suggest that the construct may not necessarily be dyadic, although others may influence students' engagement as part of their environment. Methodologically, the study's results point at the importance of careful consideration of measurement and analytic issues in researching engagement as a situational construct (see also the Special Issue on engagement in the context of science learning in Educational Psychologist, Sinatra et al., 2015). Based on the current findings, we would suggest to operationalize engagement as behavior observable from moment-to-moment in different settings. Practically, the study's results seem to suggest that teacher strategies to increase engagement would need to be extended beyond stimulating individual *students* (Yazzie-Mintz & McCormick, 2012) to include design of 'optimal' learning *situations*. Ultimately, such optimal situations could lead to classrooms full of mentally, affectively and behaviorally involved students participating in activities that create opportunities for learning.

Declaration of Competing Interest

None.

Appendix A

Code book On-task behaviour .

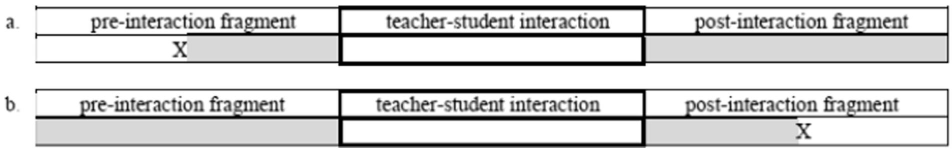

Fig. A1. Pre-interaction – interaction – post-interaction sequences entailing shorter pre- (a) and post- (b) interaction fragments. Xs represent a teacher-student interaction or whole-class remark; shaded areas depict coded parts of pre-/post-interaction fragments.

General rules

- Keep track of questions and choices you doubt about in a separate Word document; also report starting and ending times of fragments shorter than 00:02.
- Also code fragments if only a few (first or last) seconds should be coded (e.g., 2893_02); the coding will then consist of only a few coded seconds.
- IF a new behavior occurs, but this behavior takes less than 2 sec., THEN don't change codes.

Specific rules

n/f

- IF student listens to teacher's/other student's explanation,
AND this explanation is not about the student's current task, THEN code f.
AND this explanation is about the student's current task, THEN code n.
(If it is unclear, then 'benefit of the doubt': code n.)
- IF student follows teacher's behavior/listens to teacher/looks at blackboard during whole-class instruction, THEN code n (we assume that he/she participates, thinks along).
- IF student looks at (exercise) book of fellow student,
AND both students are working on the same exercise, THEN code n.
AND this student is working on a different exercise, THEN code f.
- IF student looks up
AND appears to be thinking about the task at hand, THEN code n.
AND appears to be daydreaming or distracted, THEN code f.
* Example: 296_01 01:20–01:50.
(If it is unclear, then 'benefit of the doubt': code n.)

Table A1
Codes, descriptions, and examples for coding on-task behavior.

	On-task (n)	Off-task (f)	Creating conditions (c)	Uncodable (u)
Description	A student shows on-task behavior when (s)he is attentive to and concentrated on[1] his/her *own task*, either actively (e.g., working, writing) or passively (e.g., listening)[2], or when (s)he is involved in activities expected or instructed by the teacher[3]. [1] Skinner et al., 2009 [2] cf. Volpe et al., 2005 [3] Imeraj et al., 2013	A student shows off-task behavior when his/her attention is directed away from his/her *own task for 2 or more seconds* [4], when (s)he is involved in activities not expected or instructed by the teacher[5], or when (s)he shows signs of mental withdrawal or giving up[6]. [4]Antonini, Narad, Langberg, & Epstein, 2013; Imeraj et al., 2013 [5] Imeraj et al., 2013 [6] Skinner et al., 2009	When a student is creating the conditions for on-task behavior (e.g. raising hand and waiting to ask a question, sharpening pencil), the behavior is coded as 'creating conditions'.	If student is off-screen/ inaudible/invisible AND it is not possible to derive from context and non-verbal behavior what the student is doing.
Examples	- looking away while counting to themselves or counting their fingers (Antonini, Narad, Langberg, & Epstein, 2013) - discussion about task with fellow student or teacher - student responds to task-related question of the teacher or a fellow student (both verbal and non-verbal) - listening to teacher's task-related explanation: being oriented toward the teacher or task and responding verbally (e.g., asking questions about the instructions) or nonverbally (e.g., nodding) (Allday & Pakurar, 2007) - working on the task (flipping through book is also considered on-task if it is necessary for the task at hand) - asking the teacher a task-related question	- maintaining visual attention during the task (e.g., looking at worksheet/ workbook), but at the same time doodling, counting problems, or looking ahead to see how many problems were left (Antonini, Narad, Langberg, & Epstein, 2013) - staring /looking around (Lutz et al., 2006) - listening to off-task talk of other students - listening to teacher, who is talking about something off-task - looking up when something off-task happens - talking about non-task related topics (Skinner et al., 2009) - not participating in assigned activity or not working on assigned tasks (Hart, Massetti, Fabiano, Pariseau, & Pelman Jr., 2011) - building paper airplanes or the like (Skinner et al., 2009) - daydreaming (Skinner et al., 2009) - head completely down on desk (Lutz et al., 2006) - teacher has to tell student to get to work (Lutz et al., 2006) - responds to non-task related question of the teacher or a fellow student	- raising hand - sharpening pencil - flipping through the book to find the assigned tasks - walking to teacher or to own seat - logging in onto the computer	

Table B1
Diagnostic strategy codes with description and examples.

Code	Explanation	Examples
0	No diagnostic question or a question in combination with support.	*see* Table B3
1	Diagnostic question (or follow-up of diagnostic question); either on meta level or subject matter level, either eliciting a short or elaborated response.	* That is still a bit difficult for you, isn't it? * Eva, are things ok now with the-with the verbs and that you had found the finite verb and that you would not write down a finite verb for another verb? * Which word do you mean? * What does 'to scale' mean? You learned that in geography class. * What is it that you don't understand? * What does that say, when your w is zero?

- IF fellow student asks focus student something unrelated to the specific task that the focus student is working on, THEN code f (e.g., 4907_02).
- IF student looks at task (exercise book, book, etc.), then code n (benefit of the doubt, e.g., 3270_02).

C

- IF student is flipping through his/her (exercise) book
  during work (e.g., to find the theory related to the exercises that he/she is working on), THEN code as n.
  before starting to work (e.g., to find the assigned tasks), THEN code as c.
- IF a student is cleaning up, THEN code c (e.g. 2816_01).
- IF the teacher cannot see that the student wants to ask something (e.g., when student picks up a pen that fell down while having his hand raised, 3738_01), THEN code f (not c).
- IF student turns pages while working on the task, THEN code n (not c).
- IF student looks up at the blackboard to check what tasks should be done, THEN code c (e.g., 1155_02).

U

- IF student walks off-screen, THEN code u.
- IF the student re-enters, THEN restart coding.
- IF the student re-enters AND it becomes clear that what the student has just done was c (e.g. fetching a dictionary), THEN code the off-screen part c.
- IF student is invisible but you can see/hear that another student/teacher is talking to him/her, THEN do not code u but try to determine whether it is n, f, or c.

Only code u when you have *no clue* about student's behavior (completely invisible, no clues in students around that student).

Appendix B

Coding scheme adaptive support, see (van de Pol, Mercer, & Volman, 2019; van Halem, 2016)

I. General Coding Rules
- If a turn is not audible, previous or following turns could be used as a source of information.
- Codes are assigned to a transcribed turn, but still, the interpretation of the video material is superior over the interpretation of the transcripts when assigning codes.
- Interpret as little as possible. However, gestures such as nodding, which could be translated into a word, may be translated in a yes, no, etc.

II. Coding Diagnostic strategies

Since there are different contingency rules for diagnostic questions, these have to be identified separately. The codes are presented in Table B1; original distinctions in types of diagnostic strategies used in Van Halem (2016) were collapsed.

Coding rules

- Diagnostic strategies are used to explore the level of understanding of the student. These strategies are a way to measure the abilities/knowledge of the student so far and accordingly the point at which the student needs help. Therefore, diagnostic strategies always elicit a response and the teacher waits for this response (if not, it is not a diagnostic strategy anymore, 5 s of quietness also counts as a response).
- Diagnostic strategies are applied at the start of the interaction; otherwise it would be 'checking understanding'. As soon as the teacher provides some extent of support to the student, further turns are not coded as a diagnostic strategy. Thus, diagnostic strategies are found in the first, and maybe also in some following turns (if these are follow up questions), unless: 1) The first turns do not address the question of the student (if it is not about subject matter, e.g. What are you saying?), 2) A new task/problem is addressed after the first question of the student is *resolved*.
- When there is both support and a diagnostic strategy in one turn, this diagnostic strategy doesn't count.
- A follow-up question is only coded as a diagnostic strategy when it is a follow-up question of a diagnostic strategy. In that case, the follow-up question can be a diagnostic check, such as 'so you don't know…?'.
- Diagnostic strategies can take place either on a *meta*-level (What is it you don't get.) or on the subject level (What did we discuss last time about …).
- Diagnostic strategies either refer to something closely related to the question of the student or to something that came up previously. The latter could be recognized by signal words such as: Last time, do you remember, already, what again is…, etc.
- Sometimes the teacher asks for explanations as a diagnostic strategy, such as 'Why did you choose …', 'Can you explain why…', 'What do you mean by …'.
- Reading the students work is not included in the sequences, though this is a diagnostic strategy. If it happens, this should be indicated separately, so that in retrospect it could be determined whether a kind of diagnostic strategy was used.
- When code 2 and 3 apply both to a teacher turn, code 2 counts.
- When an interaction starts with 'Yes?', 'Yes', of [Name?], this should not be coded as a diagnostic strategy. This should be regarded as a turn that doesn't addresses the question of the student yet, so following turns could be coded as diagnostic strategies.
- When a teacher quotes the task at the start of the interaction, this should be coded as a diagnostic strategy. Reading the task out loud invites the student to provide information about his/her level of understanding, though it does not provide new content.
- If the student is quiet longer than 5 s, this counts as a response.
- If the teacher asks a student to repeat what (s)he said, because the teacher didn't hear what the student said, this does not count as a diagnostic strategy.

III. Coding Teacher Support

Codes used to code the two dimensions of teacher support are presented in Table B2, including explanations and examples.

Coding rules control – form:

- When both an open question and a closed question are identified in the teacher turn, the closed one counts (in all cases). The closed question provides a level of control that can't be undone.
- Sometimes the teacher lets the student finish a sentence (e.g. the square of three is…?), depending on the type of elicitation (focused on a short or long answer), this should be coded as an open or closed question.

Coding rules control – content:

- When both level 1 and 2 apply to a teacher turn, code 2 counts, because level two provides a level of control that can't be undone.
- *New content* is (a piece of) information that is relevant to solve the problem of the student. The content could be a concept, a translation, a rule, a relation, a characteristic, or another piece of information. New content could be offered in different ways as well; it could be demonstrated, modeled, explained, instructed, etc.
- Steering is no new content per se, since it is assumed that all teacher turns provide some extent of steering. Quoting or reading from the book/other paper is no new content when it is a quote from the task, other quotes (relevant to the problem of the student) should be determined according to the other rules about new content.
- *New content* is identified in relation to the general problem that should be solved, which is mostly the help-question of the student at the start of the fragment or the particular task that the help-question addresses.
- *New content* is seldom presented in a question. (E.g. is that a positive or a negative number? When positive/negative wasn't mentioned before is no new content, though it is a steering question.) New content could show in subordinate clause of a question or another kind of addition to the question. (E.g. to announce is the predicate (new content), but what is the subject?) Suggestions are the exceptions to this rule. A suggestion that is posed in an interrogative way is coded as new content. (E.g. did you already think of the internal market? (When this concept was not mentioned before.)
- Pointing to something should not be coded.
- Repetition of something that is said previously is no new content (although it can be seen as extra support, it does not provide new information); when a turn contains both repeated and new content, it should be coded as new content.
- Feedback is seen as partial new content (code 1). When the teacher is feeding back, the teacher is providing the student information on whether he/she is doing the right thing, found the right answer, needs to do something more, etc. (E.g. 'perfect', 'good', 'that is completely off'.)
- If doubt about whether particular turn is suggestion (code 1) or instruction (code 2), only choose 2 if the instruction in the turn is not free of obligations.
- Instructions are only coded as 2 if these almost give the answer away or provide very clear guidance as to how the answer can be acquired.
- A non-completed turn that is completed in another teacher turn should be coded 0 (except when the non-completed turn already provides some useful and supportive information); the turn in which the actual support is given (completed turn) is coded following the rules of coding support (see Table B3).

IV. Coding Student Understanding

Codes used to code student understanding are presented in Table B4, including descriptions and examples. Originally, student understanding was coded using two dimensions (mode and accuracy). In the current thesis, we do not distinguish between different modes of understanding (i.e., claim or demonstration); mode of student understanding, therefore, is not included in this codebook (see Table C1).

Coding rules:

- When a turn does not show (partial (1)) *mis*understanding (0), code full understanding (2). So If a student says ''now I get it'/'I thought that it was only an area', i.e. a claim, code 2. Instead of evidence for understanding, search for evidence of misunderstanding.
- Accuracy should be determined on a micro level. So, is the student getting what the teacher says in (one of the) previous turns? Not whether or not the student is getting the whole problem/task. The latter is too complex too judge reliably.
- Repetition of what a teacher said previously is not coded differently from other turns.
- If a teacher asks a follow-up question when more info is needed according to the teacher, code 1. Only code 1 if the teacher wants more information about the response that the student has given or adds something to the response of the student. If it is a follow-up question, but the student response to the first question was correct, then code 2.
- Feedback/reaction of the teacher in one of the following turns is superior to 'common sense', since information about the task is missing for the coders. When a teacher in any way confirms the

Table B2
Overview of adaptive support dimensions and possible codes for coding three-turn sequences.

Codes	Teacher turns			Student turns
	Diagnostic Strategy	Control – form (coded if turn is not a Diagnostic Strategy)	Control – content (coded if turn is not a Diagnostic Strategy)	Accuracy of Student Understanding
0	No diagnostic question or a question in combination with support.	Not formatted as a question.	Support contains no new content (though it can provide a direction, through a steering question).	No understanding (no answer, incorrect answer, or guessing).
1	Diagnostic question (or follow-up question), exploring student's level of understanding.	Formatted as a closed question (elicits a response that *can* be reduced to one word or concept).	Support contains partial new content, through explicit clues, suggestions, tips, or feedback.	Partial understanding (partial correct information/ information is missing).
2	*not applicable*	Formatted as an open question (elicits a response that cannot be summarized in one word or concept).	Support provides direct new content, i.e., instructions or explanations, or provides the opportunity for imitation.	Full understanding (provides correct answer, shows understanding).

Note. Adapted from "Reconsidering Variability in Support Adaptivity in the Classroom: The Unique Contribution of Micro and Dyad Level Factors to Support Adaptivity" (Van Halem, 2016).

Table B3
Teacher support codes with description and examples.

Code	Control – form Description	Examples	Control – content Description	Examples
0	No question	* The answer is prosperity * Twenty-four * This one. * Subtract 2.	Support contains no new content. (Though it can provide a direction, through a steering question).	* Which concept can be left out? * Who are 'they'? (follow up question)
1	Closed question (elicits a response that *can* be reduced to one word or concept)	* Do you think so? * Yes?/ No? * Which concept can be left out? * Who did that, the Catholics or the Protestants? * What's his name?/How is that called? * The square of 10 is… ?	Support contains partial new content, through explicit clues, suggestions, tips, or feedback.	* Have you thought about the internal market [when this concepts wasn't mentioned yet] * It is depicted somewhere on the page. * This argument is wrong * I've only heard you saying what import duty is, it is some kind of duty that has to be paid so that's an explanation of the concept but I haven't heard the reason why import duty can be left out of the series.
2	Open question (elicits a response that can't be summarized in one word or concept)	* Why did you categorize this newspaper article into category X? * What do you think that prosperity means? * Why do you think it should be rights? * Can you explain that? * What is that protuberance?	Support provides direct new content, i.e. instructions, explanations, or pro-vides the opportunity for imitation.	* The answer is prosperity * S: So this article goes with prosperity? T: yes * Twenty-four (as the outcome of a formula) * S: Which one is wrong? T: This one. * Subtract 2. (Direct instruction)

Table B4
Student understanding codes with description and examples.

Code	Explanation	Examples
0	No understanding (No answer, incorrect answer, or guessing)	* T: do you get it? S: no * S: Could it also just be this school? T: No, because people don't visit this place, do they? * Is the West actually above sea level? T: What do you think? S: Yes? (guessing) * S: Ligt het westen eigenlijk boven zeeniveau? T: Wat denk je? S: Ja? (guessing)
1	Partial understanding (partial correct information/partial incorrect or information is missing)	* T: What do these *three* concepts have in common, border – customs – internal market? S: well, the customs are at the border (SA2) [the student only refers to two concepts] * S: How is that thing with kilometers called? ? T: A cubic kilometer. S: Cubic kilometer. * T: What is the protrusion? S: This leaf. T: The protrusion is the leaf and the root.
2	Full understanding (provides the correct answer/shows understanding)	* S: I know this one.; T: Do you understand? S: Yes. * T: To which category does this article belong? S: environment. T: perfect. * T: "Farmers get subsidised by the EU to buy new goats" Does that go with prosperity? S: yes. * T: what does prosperity mean? S: How well a country is doing. T: yes, it is about spending money

student's answer, the student's answer should be coded as full understanding.

- If a teacher initially says that a student response is not correct, while it appears to be correct later in the interaction (according to the teacher), then code the student response as 2.
- If a student asks the teacher to come and asks a question (no matter whether the question is finished or not), then code 2.
- If a student says o/oh/ok/yes, the context determines which code is assigned:

  If the teacher asks for a claim of understanding, then o/oh/ok/yes can be 2.

  If o/oh/ok/yes only means 'yes, continue/yes, I've heard you', then code 999.

  If the o/oh/ok/yes is not preceded by a teacher elicitation of claim of understanding, then the intonation of o/oh/ok/yes determines the code. A clear oooooohhh! or o yes! should be coded 2, otherwise 999.
- Code 0 is assigned when:

A. The students are quiet for a long time (more than 5 s) after a teacher question, or utters just some words (e.g., 'hmmm') without giving an answer,
B. When student or teacher claim not knowing or not understanding,
C. A student requests explanation or support/asks a question. Providing the good answer with a 'questioning' voice is the exception to this, because this is full or partial understanding. If the student response in questioning voice is correct, then code 2
D. If the student says something that does not apply/is not the point or correct
E. The student is guessing or joking.

- Code 999 if unintelligible or not related to understanding.

Appendix C

Coded example of Adaptive support

Table C1
Example of coded interaction (codings and scores) between student 3408 and teacher 18.

Turn	Codings				Contingency scores			
	DS	C1	C2	SU	C_f (sequence)	C_c (sequence)	C1 + C2 (sequence)	Aggregated (interaction)
ST3408 Ehm.. This is one sentence part [zinsdeel], isn't it?				0				0.88
TCHR18 Is it?	1	1	0					
ST3408 Well, that's what I thought…				0	0	0	0	
TCHR18 What is, what is the predicate?	0	1	0					
ST3408 Ehm… come? back?				2	1	0	1	
TCHR18 Yes, well done!	0	0	1					
ST3408 Oh.				999	999	999		
TCHR18 And then there is still something left?	0	1	0					
ST3408 Immediately.				2	1	0	1	
TCHR18 Yes!	0	0	1					
ST3408 And how exactly should I mark it?				0	1	1	2	
TCHR18 Yes so then you should (points) here a line and there a line. So two of them. You see?	0	1	2					
ST3408 Yes.				2	1	0	1	
TCHR18 Because immediately, you know, would be a verb i fit belonged to the predicate and is isn't a verb.	0	0	2					
ST3408 Ok. Is back a verb too?				0	1	0	1	
TCHR18 Which one?	0	1	0					
ST3408 (points)				0	0	1	1	
TCHR18 No but no, that goes together, back alone is not a verb of course, but together it is. Just like out, out actually is a preposition.	0	0	2					
ST3408 So then it actually is a object or something?				0	0	0	0	
TCHR18 (laughs) Object. No now you confuse everything.	0	0	1					
ST3408 Yes.				999	999	999		
TCHR18 This is a separable compound verb, come back, it only is separated in the sentence so these together form the predicate. This doesn't have anything to do with the object. And this is the second part sentence part [zinsdeel].	0	0	2					
ST3408 Ok.				2				

Note. DS = diagnostic strategy; C_f = control – form; C_c = control – content; SU = student understanding.

Appendix D. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cedpsych.2020.101928.

References

Albright, J. J., & Marinova, D. M. (2010). Estimating multilevel models using SPSS, Stata, and SAS. Indiana University; 2010. Retrieved fromhttp://www.iub.edu/%7Estatmath/stat/all/hlm/hlm.pdf.

Anderson, L. (1984). The environment of instruction: The function of seatwork in a commercially developed curriculum. Comprehension Instruction: Perspectives and Suggestions, 93–103.

Allday, R. A., & Pakurar, K. (2007). Effects of teacher greetings on student on-task behavior. *Journal of Applied Behavior Analysis, 40*(2), 317–320.

Antonini, T. N., Narad, M. E., Langberg, J. M., & Epstein, J. N. (2013). Behavioral correlates of reaction time variability in children with and without ADHD. *Neuropsychology, 27*(2), 201–209.

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*(5), 369–386. https://doi.org/10.1002/pits.20303.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*(5), 427–445. https://doi.org/10.1016/j.jsp.2006.04.002.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist, 50*(1), 84–94. https://doi.org/10.1080/00461520.2015.1004069.

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences.* Cambridge: Cambridge University Press.

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology, 35*(1), 61–79.

Briesch, A. M., Chafouleas, S. M., & Riley-Tillman, T. C. (2010). Generalizability and dependability of behavior assessment methods to estimate academic engagement: A comparison of systematic direct observation and direct behavior rating. *School Psychology Review, 39*(3).

Cadima, J., Doumen, S., Verschueren, K., & Buyse, E. (2015). Child engagement in the transition to school: Contributions of self-regulation, teacher–child relationships and classroom climate. *Early Childhood Research Quarterly, 32*, 1–12. https://doi.org/10.1016/j.ecresq.2015.01.008.

Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperative learning: How to improve student problem solving and time on-task. *American Educational Research Journal, 41*(2), 365–399. https://doi.org/10.3102/00028312041002365.

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement.* New York: Springer Science.

Eccles, J., & Wang, M.-T. (2012). Part I commentary: So what is student engagement anyway? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 133–145). New York: Springer Science.

Fielding, A., & Goldstein, H. (2006). Cross-classified and multiple membership structures in multilevel models: An introduction and review (791). Retrieved from http://dera.ioe.ac.uk/6469/1/RR791.pdf.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95*, 148–162.

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97–131). New York: Springer Science.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059.

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction, 43*, 1–4. https://doi.org/10.1016/j.learninstruc.2016.02.002.

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–782). New York: Springer Science.

Gilbert, J., Petscher, Y., Compton, D. L., & Schatschneider, C. (2016). Consequences of misspecifying levels of variance in cross-classified longitudinal data structures. *Frontiers in Psychology, 7*, 695. https://doi.org/10.3389/fpsyg.2016.00695.

Guo, Y., Sun, S., Breit-Smith, A., Morrison, F. J., & Connor, C. M. (2015). Behavioral engagement and reading achievement in elementary-school-age children: A longitudinal cross-lagged analysis. *Journal of Educational Psychology, 107*(2), 332–347. https://doi.org/10.1037/a0037638.

Gwet, K. L. (2014). Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters. *Advanced Analytics LLC.*

Hart, K. C., Massetti, G. M., Fabiano, G. A., Pariseau, M. E., & Pelman, W. E., Jr. (2011). Impact of group size on classroom on-task behavior and work productivity in children with ADHD. *Journal of Emotional and Behavioral Disorders, 19*(1), 55–64.

Hiebert, J. (2003). What research says about the NCTM standards. *A Research Companion to Principles and Standards for School Mathematics* (pp. 5–26). Reston, Va: National Council of Teachers of Mathematics.

Helmke, A., & Schrader, F.-W. (1988). Successful student practice during seatwork: Efficient management and active supervision not enough. *The Journal of Educational Research, 82*(2), 70–76. https://doi.org/10.1080/00220671.1988.10885870.

Imeraj, L., Antrop, I., Sonuga-Barke, E., Deboutte, D., Deschepper, E., Bal, S., & Roeyers, H. (2013). The impact of instructional context on classroom on-task behavior: A matched comparison of children with ADHD and non-ADHD classmates. *Journal of School Psychology, 51*(4), 487–498. https://doi.org/10.1016/j.jsp.2013.05.004.

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*(3), 588–600. https://doi.org/10.1037/a0019682.

Kenny, D. (2007). Estimation of the SRM using specialized software. Retrieved from davidakenny. net/doc/srmsoftware.doc.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford press.

Kenny, D. A., Mohr, C. D., & Levesque, M. J. (2001). A social relations variance partitioning of dyadic behavior. *Psychological Bulletin, 127*(1), 128–141. https://doi.org/10.1037/0033-2909.127.1.128.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 159–174*. https://doi.org/10.2307/2529310.

Lee, O., & Anderson, C. W. (1993). Task engagement and conceptual change in middle school science classrooms. *American Educational Research Journal, 30*(3), 585–610. https://doi.org/10.3102/00028312030003585.

Leon, J., Medina-Garrido, E., & Núñez, J. L. (2017). Teaching quality in math class: The development of a scale and the analysis of its relationship with engagement and achievement. *Frontiers in Psychology, 8*, 895. https://doi.org/10.3389/fpsyg.2017.00895.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology, 34*(2), 120–131. https://doi.org/10.1016/j.cedpsych.2008.12.001.

Lutz, S. L., Guthrie, J. T., & Davis, M. H. (2006). Scaffolding for engagement in elementary school reading instruction. *The Journal of Educational Research, 100*(1), 3–20. https://doi.org/10.3200/JOER.100.1.3-20.

Mahatmya, D., Lohman, B. J., Matjasko, J. L., & Farb, A. F. (2012). Engagement across developmental periods. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 45–63). New York: Springer Science.

Martin, A. J. (2011). *The motivation and engagement scale* (11th). Sydney, NSW: Lifelong Achievement Group.

Mainhard, T., Oudman, S., Hornstra, L., Bosker, R. J., & Goetz, T. (2018). Student emotions in class: The relative importance of teachers and their interpersonal relations with students. *Learning and Instruction, 53*, 109–119. https://doi.org/10.1016/j.learninstruc.2017.07.011.

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal, 37*(1), 153–184. https://doi.org/10.3102/00028312037001153.

Martin, A. J., & Collie, R. J. (2019). Teacher–student relationships and students' engagement in high school: Does the number of negative and positive relationships with teachers matter? *Journal of Educational Psychology, 111*(5), 861. https://doi.org/10.1037/edu0000317.

Martin, A. J., Mansour, M., & Malmberg, L.-E. (2019). What factors influence students' real-time motivation and engagement? An experience sampling study of high school students using mobile technology. *Educational Psychology, 1–23*. https://doi.org/10.1080/01443410.2018.1545997.

Martin, A. J., Papworth, B., Ginns, P., Malmberg, L.-E., Collie, R. J., & Calvo, R. A. (2015). Real-time motivation and engagement during a month at school: Every moment of every day for every student matters. *Learning and Individual Differences, 38*, 26–35. https://doi.org/10.1016/j.lindif.2015.01.014.

McKellar, S. E., Cortina, K. S., & Ryan, A. M. (2019). Teaching practices and student engagement in early adolescence: A longitudinal study using the classroom assessment scoring system. *Teaching and Teacher Education, 102936*. https://doi.org/10.1016/j.tate.2019.102936.

Motti-Stefanidi, F., Masten, A., & Asendorpf, J. B. (2015). School engagement trajectories of immigrant youth: Risks and longitudinal interplay with academic success. *International journal of behavioral development, 39*(1), 32–42.

Nie, Y., & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology, 34*(3), 185–194. https://doi.org/10.1016/j.cedpsych.2009.03.001.

Patall, E. A., Vasquez, A. C., Steingut, R. R., Trimble, S. S., & Pituch, K. A. (2016). Daily interest, engagement, and autonomy support in the high school science classroom. *Contemporary Educational Psychology, 46*, 180–194. https://doi.org/10.1016/j.cedpsych.2016.06.002.

Peterson, P. L., Swing, S. R., Stark, K. D., & Waas, G. A. (1984). Students' cognitions and time on task during mathematics instruction. *American Educational Research Journal, 21*(3), 487–515. https://doi.org/10.3102/00028312021003487.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365–386). New York: Springer.

Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2018). Variation in situation-specific engagement among lower secondary school students. *Learning and Instruction, 53*, 64–73. https://doi.org/10.1016/j.learninstruc.2017.07.007.

Reschly, A. L., & Christenson, S. L. (2006). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education, 27*(5), 276–292. https://doi.org/10.1177/07419325060270050301.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*(3), 700–712. https://doi.org/10.1037/a0027268.

Ruzek, E. A., & Schenke, K. (2019). The tenuous link between classroom perceptions and motivation: A within-person longitudinal study. *Journal of Educational Psychology, 111*(5), 903–917. https://doi.org/10.1037/edu0000323.

Santoyo, C., Jonsson, G. K., Anguera Argilaga, M. T., & López Lopez, J. A. (2017). Observational analysis of the organization of on-task behavior in the classroom using complementary data analyses. *Anales de Psicología, 33*(3), 497–514.

Shernoff, D. J., Kelly, S., Tonks, S. M., Anderson, B., Cavanagh, R. F., Sinha, S., & Abdi, B. (2016). Student engagement as a function of environmental complexity in high school classrooms. *Learning and Instruction, 43*, 52–60. https://doi.org/10.1016/j.learninstruc.2015.12.003.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924.

Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks: Sage.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493–525. https://doi.org/10.1177/0013164408323233.

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21–44). New York: Springer Science.

Suárez-Orozco, C., Suárez-Orozco, M. M., & Todorova, I. (2009). *Learning a new land: Immigrant students in American society*. Cambridge: Harvard University Press.

Tropper, N., Leiss, D., & Hänze, M. (2015). Teachers' temporary support and worked-out examples as elements of scaffolding in mathematical modeling. *ZDM, 47*(7), 1225–1240.

Taboada Barber, A. M., Buehl, M. M., & Beck, J. S. (2017). Dynamics of engagement and disaffection in a social studies classroom context. *Psychology in the Schools, 54*(7), 736–755. https://doi.org/10.1002/pits.22027.

van Halem, N. (2016). Reconsidering Variability in Support Adaptivity in the Classroom: The Unique Contribution of Micro and Dyad Level Factors to Support Adaptivity [Master thesis]. Utrecht University.

van de Pol, J., & Elbers, E. (2013). Scaffolding student learning: A micro-analysis of teacher–student interaction. *Learning, Culture and Social Interaction, 2*(1), 32–41. https://doi.org/10.1016/j.lcsi.2012.12.001.

van de Pol, J., Mercer, N., & Volman, M. (2019). Scaffolding student understanding in small-group work: Students' uptake of teacher support in subsequent small-group interaction. *Journal of the Learning Sciences, 28*(2), 206–239. https://doi.org/10.1080/10508406.2018.1522258.

van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2014). Teacher scaffolding in small-group work: An intervention study. *Journal of the Learning Sciences*, 600–650.

van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: Support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science, 43*(5), 615–641. https://doi.org/10.1007/s11251-015-9351-z.

Vasalampi, K., Muotka, J., Pöysä, S., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2016). Assessment of students' situation-specific classroom engagement by an InSitu Instrument. *Learning and Individual Differences, 52*, 46–52. https://doi.org/10.1016/j.lindif.2016.10.009.

Vollet, J. W., Kindermann, T. A., & Skinner, E. A. (2017). In peer matters, teachers matter: Peer group influences on students' engagement depend on teacher involvement. *Journal of Educational Psychology, 109*(5), 635. https://doi.org/10.1037/edu0000172.

Volpe, R. J., DiPerna, J. C., Hintze, J. M., & Shapiro, E. S. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review, 34*(4), 454.

Wellborn, J. G. (1991). Engaged and disaffected action: The conceptualization and measurement of motivation in the academic domain. University of Rochester, New York.

Wood, D., Wood, H., & Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. *International Journal of Behavioral Development, 1*(2), 131–147. https://doi.org/10.1177/016502547800100203.

Yazzie-Mintz, E., & McCormick, K. (2012). Finding the humanity in the data: Understanding, measuring, and strengthening student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 743–761). New York: Springer Science.