Psychology of Sport & Exercise 60 (2022) 102174

Contents lists available at ScienceDirect

Psychology of Sport & Exercise

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

Intervention-enabled autonomy-supportive teaching improves the PE classroom climate to reduce antisocial behavior☆

Sung Hyeon Cheon [a,**], Johnmarshall Reeve [b,*], Herbert W. Marsh [b], Yong-Gwan Song [c]

[a] Korea University, South Korea
[b] Australian Catholic University, Australia
[c] Pukyong National University, South Korea

ARTICLE INFO

Keywords:
Antisocial behavior
Autonomy support
Classroom climate
Doubly latent
Self-determination theory

ABSTRACT

Background: Autonomy-supportive teaching interventions enhance PE student outcomes. According to previous research, these benefits occur because autonomy-supportive teaching enhances students' psychological needs, though they may also occur because such teaching enhances the classroom climate. The student benefit of interest was reduced classroom-wide antisocial behavior.
Objectives: We predicted that teacher participation in the intervention would enhance both classroom climate and psychological needs assessed at the classroom level. We further predicted that improvements in the classroom climate would better explain decreased antisocial behavior.
Method: Using a cluster randomized control trial design with longitudinally-assessed dependent measures, we randomly assigned 49 physical education secondary-grade Korean teachers to participate (or not) in an autonomy-supportive teaching intervention (25 experimental, 24 control). The 1487 students in these 49 classrooms reported their individually-experienced need satisfaction and frustration and their classroom-level supportive climate, conflictual climate, and antisocial behavior across three waves.
Results: A series of doubly latent multilevel structural equation modeling analyses showed that, at the classroom level, (1) intervention-enabled autonomy-supportive teaching improved both students' psychological needs (more satisfaction, $\beta = 0.84$; less frustration, $\beta = -0.66$) and the prevailing classroom climate (more supportive, $\beta = 0.77$; less conflictual, $\beta = -0.68$) and (2) the improved climate best explained why antisocial behavior declined (overall $R^2 = 0.86$).
Conclusion: These findings show the importance of incorporating classroom climate effects to understand why autonomy-supportive teaching interventions improve student outcomes.

Randomized control trials confirm that teacher participation in an autonomy-supportive teaching intervention produces many student benefits (Reeve & Cheon, 2021). Students of physical education (PE) teachers who have been randomly assigned to participate in a carefully-designed, theory-based workshop on how to become more autonomy supportive (experimental condition) display all of the following end-of-semester gains in their academic, personal, and social functioning (compared to students of teachers in a "practice as usual" control condition):

- Greater classroom engagement (Cheon et al., 2019; Cheon et al., 2016);
- Greater agency and initiative and lesser passivity (Reeve, Cheon, & Yu, 2020);
- Lower sedentary level (Lonsdale et al., 2013);
- Greater task absorption (Ulstad, Halvari, Sorebo, & Deci, 2018);
- More positive, less negative emotions (Flunger, Mayer, & Umbach, 2019; Kaplan & Assor, 2012);

☆ This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5A2A01041790) and by Korea University Grant.
* Corresponding author. Institute for Positive Psychology and Education, Australian Catholic University, North Sydney Campus, 33 Berry Street, 9th floor, Sydney, Australia, 2060.
** Corresponding author. Department of Physical Education, Korea University, Anam-dong, Seongbuk-gu, Seoul, 02841, South Korea.
E-mail addresses: cheon78@korea.ac.kr (S.H. Cheon), johnmarshall.reeve@acu.edu.au (J. Reeve).

https://doi.org/10.1016/j.psychsport.2022.102174
Received 3 September 2021; Received in revised form 2 March 2022; Accepted 2 March 2022
Available online 4 March 2022
1469-0292/

- Greater course-specific skill development (Cheon, Reeve, & Vansteenkiste, 2021; Manninen, Deng, Hwang, Waller, & Yli-Piipari, 2020);
- Lesser problematic relationships (Cheon, Reeve, & Song, 2019; Cheon et al., 2021);
- Lesser burnout (Langan, Toner, Blake, & Lonsdale, 2015);
- Lesser acceptance of cheating as okay (Cheon, Reeve, & Ntoumanis, 2018);
- More positive self-concept (Cheon, Reeve, & Song, 2019); and
- Greater academic achievement and improved course grades (Cheon, Reeve, & Moon, 2012, 2021; deCharms, 1976; Ulstad et al., 2018).

Why do these student benefits occur? The above investigations adopted a self-determination theory framework (SDT; Ryan & Deci, 2017) to test and empirically support a needs-based mediation model. In this explanatory model, teacher participation in the intervention enhances students' need states, and these satisfied need states then explain the corresponding gains in the course-related outcomes (Abula et al., 2020; Cheon & Reeve, 2013, 2015; Cheon et al., 2016; Cheon et al., 2019; Tilga, Hein, & Koka, 2019), thereby confirming the explanatory power of SDT's needs-based mediation model (Cheon et al., 2012; Cheon et al., 2016, 2018, 2019).

As SDT researchers expanded their attention to include negative as well as positive student outcomes, they adopted the dual-process model (Bartholomew et al, 2011a). This model proposes that adaptive outcomes (e.g., prosocial behavior) are best explained by adaptive environmental (autonomy support) and personal (need satisfaction) events, while maladaptive outcomes (e.g., antisocial behavior) are best explained by maladaptive environmental (interpersonal control) and personal (need frustration) events (Gunnell, Crocker, Wilson, Mack, & Zumbo, 2013; Haerens, Aelterman, Vansteenkiste, Soenens, & Van Petegem, 2015). The present study adopted this dual-process model framework to focus on the outcome of reduced antisocial behavior. While this framework could easily focus on promoting prosocial behavior, we focused on reducing antisocial behavior for two reasons. First, those in both sports (Kavussanu, 2012; Shields, Funk, & Bredemeier, 2015) and education (Juvonen & Graham, 2014) embrace an urgency (a societal necessity) to understand and reduce antisocial behavior. Second, we wanted to capitalize on previous SDT-based intervention research that laid the groundwork to help us understand how and why greater autonomy-supportive teaching reduces antisocial behavior (Assor, Feinberg, Kanat-Maymon, & Kaplan, 2018; Kaplan & Assor, 2012; Roth, Kanat-Maymon, & Bibi, 2010). These researchers showed that autonomy-supportive teachers reduced classroom violence by promoting caring (Assor et al., 2018), by listening to and accepting students' concerns (Kaplan & Assor, 2012), and by helping students volitionally internalize teacher recommendations (Roth et al., 2010).

1. The classroom as the unit-of-analysis

To test for the benefits of an autonomy-supportive teaching intervention, researchers routinely use the individual student as the unit-of-analysis and investigate only for individually-experienced pathways (i.e., need satisfaction) to student outcomes. This exclusive focus on the individual student as the unit-of-analysis is an important limitation for three reasons.

First, researchers offer autonomy-supportive teaching interventions to the teacher—not to individual students. The classroom is actually the better and more appropriate unit-of-analysis in teacher intervention studies. This is because classroom observational studies show that global teachers, on average, spend about 88% of class time in front of the class providing whole-class instruction—that is, one teacher provides the same instruction simultaneously to 30 or so students (OECD, 2020). Teachers do sometimes provide individualized instruction, but mostly teachers provide group instruction.

Second, autonomy-supportive teaching enhances not only the quality of students' psychological needs but also the quality of peer-to-peer interactions and relationships. In a pair of studies (Assor et al., 2018; Kaplan & Assor, 2012), researchers used an autonomy-supportive teaching intervention to help teachers initiate teacher-to-whole-class dialogues both to create a more supportive peer climate and to decrease antisocial behavior (though these studies also used only the individual student as the unit-of-analysis).

Third, antisocial behavior likely emerges from some combination of individually-experienced need states and group-based social interactions and relationships (Cook, Williams, Guerra, Kim, & Sadek, 2010). Educators generally recognize antisocial behavior as a community-generated and community-regulated behavior (Hendrickx, Mainhard, Boor-Klip, Cillessen, & Brekelmans, 2016). This is because being surrounded by conflictual peer-to-peer interactions and relationships promotes antisocial behavior, while being surrounded by supportive peer-to-peer interactions and relationships prevents it (Karna et al., 2011; Kaplan & Assor, 2012; Roth et al., 2010). Antisocial behavior tends to be especially high when students find themselves in classrooms that are conflictual (Van Petegem, Soenens, Vansteenkiste, & Beyers, 2015), competitive ("me vs. you") and intimidating (Pellegrini, 2001), hierarchical (Espelage, Holt, & Henkel, 2003; Thomas and Bierman, 2006), and status-centric (Lansford, Malone, Dodge, Pettit, & Bates, 2010; Mikami, Lerner, & Lun, 2010).

Together, these three limitations suggest the need to expand a student unit-of-analysis in autonomy-supportive teaching intervention research to incorporate a classroom unit-of-analysis. We, therefore, introduce the "peer ecology" (Hendrickx et al., 2016) or the "classroom climate" (Hodge & Gucciardi, 2015; Marsh et al., 2011; Ntoumanis, Vazou, & Duda, 2007) as a potentially important group-mediated pathway to lesser antisocial behavior.

2. Classroom climate

Classroom climate is a social-ecological concept (Espelage & Swearer, 2011; Hong & Espelage, 2012) that represents the group consensus as to what behaviors are acceptable and normative. Once established, this group consensus guides the quality of the peer-to-peer interactions that occur in that classroom (Thornberg, Wanstrom, & Jungert, 2018). Different dimensions can represent the classroom climate. However, the distinction between *hierarchical and conflictual* vs. *egalitarian and supportive* is a key classroom climate dimension relevant to understanding antisocial behavior, because antisocial behavior is often a status enhancement strategy students use to advance themselves into a socially dominant position within the peer group (Salmivalli, 2010; Roth et al, 2010; Salmivalli, 2010).

A hierarchical and conflictual classroom climate arises from student-to-student (peer-to-peer) interactions and relationships that promote a status-centric dominance hierarchy as well as the norms, expectations, values, group dynamics, social roles, and patterns of communication that create and maintain a prevailing interpersonal tone of conflict and competition (Hodge & Gucciardi, 2015; Ntoumanis et al., 2007). To represent such a conflictual climate, we adopted the *peer ego-involving climate* from the sports literature that emphasizes interpersonal competition, social comparisons, and normative ability hierarchies (Joesaar, Hein, & Hagger, 2012; Ntoumanis & Vazou, 2005; Vazou et al., 2006). However, in the present study, we modified this label to downplay its motivational connotation ("ego-involvement"; Boardley & Kavussanu, 2010) in favor of its salient relationship dynamic ("conflictual"). We expected that if greater autonomy-supportive and lesser controlling teaching could diminish such a conflictual climate, then antisocial behavior in that classroom would decrease proportionally (Cheon et al, 2019; Ntoumanis et al., 2007).

An egalitarian and supportive classroom climate arises from student-to-student interactions and relationships that promote an equal hierarchy as well as the norms, expectations, values, group dynamics, and patterns of communication that create and maintain a prevailing interpersonal tone of acceptance, support, and community (Hodge & Gucciardi, 2015; Morin, Marsh, Nagengast, & Scalas, 2014; Ntoumanis et al., 2007). To represent such a supportive climate, we adopted the *peer task-involving climate* from the sports literature that emphasizes accepting and including one's peers, verbally encouraging classmates, and working together for improvement and task mastery (Joesaar et al., 2012; Ntoumanis & Vazou, 2005; Vazou, Ntoumanis, & Duda, 2006). We again modified this label from the sports literature to downplay its motivational connotation ("task-involvement") in favor of its salient relationship dynamic ("supportive"). We expected that if greater autonomy-supportive and lesser controlling teaching could encourage such a supportive climate, then that might provide a second classroom catalyst to decrease antisocial behavior (Cheon, Reeve, & Ntoumanis, 2019; Ntoumanis et al., 2007).

3. Why would autonomy-supportive teaching improve the classroom climate?

Autonomy-supportive teaching is the adoption of a student-focused attitude and an understanding tone that enables the skillful enactment of seven autonomy-satisfying instructional behaviors (e.g., take the students' perspective, present learning activities in need-satisfying ways) (Patall et al., 2018; Reeve, Ryan, Cheon, Matos, & Kaplan, 2022). An autonomy-supportive teaching intervention provides teachers with a professional development opportunity to learn how to upgrade the quality of their classroom motivating style (i.e., more autonomy-supportive, less controlling; Reeve & Cheon, 2021). During such an intervention, teachers learn how to (1) enact autonomy-supportive instructional behaviors and (2) transform existing controlling instructional behaviors into autonomy-supportive alternatives (e.g., replace "utter directives" with "provide explanatory rationales for teacher requests").

The reason why we expected autonomy-supportive teachers (ASTs) to foster a more supportive peer-to-peer climate is because ASTs begin instruction by taking their students' perspective. By doing so, ASTs model and value empathy. ASTs further adopt an interpersonal tone of understanding during instruction. In doing so, ASTs let their students know that they care about how they are feeling; are listening to students; are paying attention to their concerns; are "on their side"; are working to understand why students might be fussing and complaining; and are working to find ways to adjust instruction so it better aligns with what students need and prefer (Kaplan & Assor, 2012; Reeve et al., 2022). When ASTs make a request ("work together") or recommend a behavior ("use respectful language"), they explain the personal benefits students can expect from acting in such supportive ways (e.g., stronger friendships, more collaboration, improved performance). Together, through their perspective-taking, empathic concern, and explanatory rationales for peer support, ASTs promote a supportive peer-to-peer classroom climate.

Similarly, we expected ASTs to diminish a conflictual peer-to-peer climate because of how ASTs handle classroom discipline, interpersonal conflicts, and behavior change requests. ASTs address these classroom situations by first taking the student's perspective ("How do you feel about the language other students use?") and communicating an interpersonal tone of understanding ("I would like to understand why you behave in this way."). Instead of uttering a directive or command to resolve the situation, ASTs acknowledge the potential validity of students' negative feelings ("I understand why you feel this way; I suspect that many of your classmates may also feel the same way."), provide rationales to explain the personal value of a different way of acting, rely on invitational rather than pressuring language ("you might want to try …"), and display patience as students work through an internalization process to adjust their behavior from something that is ineffective and irresponsible to something that is more effective and more responsible. As ASTs help students revolve issues of conflict in ways that are fair, constructive, and beneficial to all parties, ASTs diminish a conflictual peer-to-peer classroom climate.

4. Hypothesized model

The purpose of the present study was to test the hypothesis that intervention-enabled improvements in the classroom climate would reduce class-wide antisocial behavior. Such an explanatory model requires a multilevel structure and analysis (Marsh et al., 2011; Morin et al., 2014), such as a doubly latent multilevel structural equation model (DL-ML-SEM; Marsh et al., 2009; March et al, 2011; Morin et al., 2014, 2021). In the present study, we created two DL-ML-SEM models: (1) a "Needs Only" model (see Figure 1) and a "Needs and Climates" model (see Figure 2). In the center of both Figures 1 and 2 is a series of three "Observed Variables:" boxes (one box for each wave of data). These horizontal bars represent students' responses to the dependent measures included on the study questionnaire. These data are used as indicators to create latent variables at both the L1 and L2 levels (hence the name "doubly latent"). The L1 latent variables are similar to those created in a traditional SEM analysis using a student unit-of-analysis.[1] The L2 latent variables are newly-added constructs to represent the classroom as the unit-of-analysis. At the L2 level, student ratings are aggregated to represent the whole-class's shared perception of a dependent measure. For both models, our hypotheses pertained to the classroom (L2) level.

Needs Only Model (Figure 1). In the Needs Only Model, we predicted that teacher participation in the autonomy-supportive teaching intervention (COND, or experimental condition) would increase students' L2 T2 need satisfaction (H1) and decrease students' L2 T2 need frustration (H2). Thus, the Needs Only Model predicted that intervention-enabled declines in L2 need frustration would explain end-of-semester declines in L2 antisocial behavior (H3). This Needs Only Model represents a mediation model, so we performed a follow-up mediation analysis. Further, in a supplemental analysis, we included the cross-over path in which low L2 T2 need satisfaction (in addition to high L2 T2 need frustration) functioned as a second supplemental predictor of L2 T3 antisocial behavior. This was not a predicted effect, but we included it to provide a comprehensive test of the dual-process model.

Needs and Climates Model (Figure 2). In the Needs and Climates Model, we predicted that teacher participation in the autonomy-supportive teaching intervention (COND) would increase both L2 T2 need satisfaction (H1) and L2 T2 supportive classroom climate (H4) and decrease both L2 T2 need frustration (H2) and L2 T2 conflictual classroom climate (H5). Following the dual-process model, the model predicted that the individually significant predictors of decrease L2 T3 antisocial behavior would be low L2 T2 need frustration (H3) and low L2 T2 conflictual climate (H6), but not necessarily greater L2 T2 need satisfaction or greater L2 T2 supportive climate. The Needs and Climates Model also represents a mediation model, so we again performed a follow-up mediation analysis. The two hypothesized mediators were L2 T2 need frustration and L2 T2 conflictual climate. Further, in a supplemental analysis, we included the two cross-over paths in which low L2 T2 need satisfaction (in addition to high L2 T2 need frustration) and low L2 T2 supportive climate (in addition to high L2 T2 conflictual climate) functioned as additional supplemental predictors of L2 T3 antisocial behavior. Again, these were not predicted effects, but we

[1] This statement holds true for latent variables created from individually-referenced questionnaires (e.g., "I feel competent"), but the meaning of an L1 latent variable created from classroom-referenced questionnaires is a little different. For the latter, the aggregated L2 variable represents the group reality (i.e., "Everyone agrees that students in this class fight.") while the corresponding L1 variable represents within-class student-to-student differences in their perception of that group reality (the residual L1 variable). We discuss the importance of this distinction in the Discussion section.

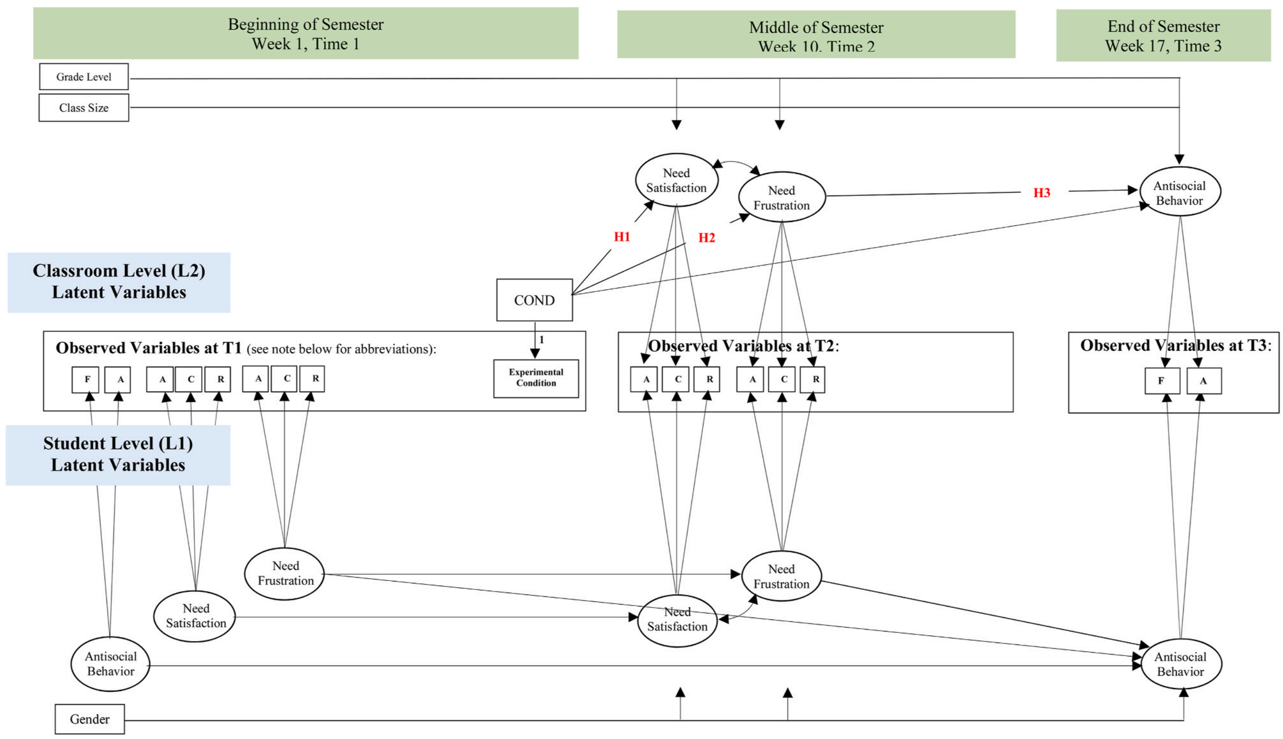

Figure 1. *Needs Only DL-ML-SEM model to Predict T3 Antisocial BehaviorNote*. Thick lines represent hypothesized paths, while thin lines represent autoregressive effects, statistical controls, or indicators of a latent variable. For Antisocial Behavior: A = Argue subscale, F = Fight subscale; For Need Satisfaction: A = Autonomy scale, C = Competence scale, R = Relatedness scale; For Need Frustration: A = Autonomy subscale, C = Competence subscale, R = Relatedness subscale; COND = Experimental condition.

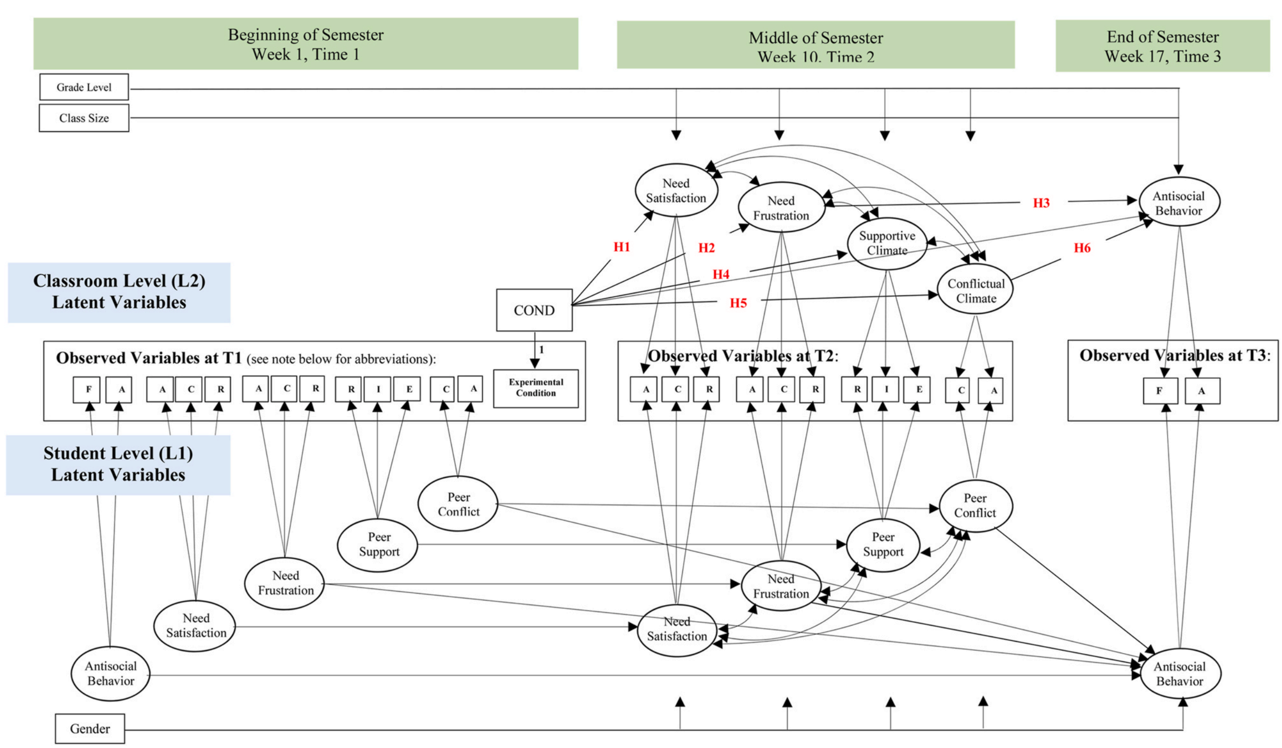

Figure 2. *Needs and Climates DL-ML-SEM model to Predict T3 Antisocial BehaviorNote*. Thick lines represent hypothesized paths, while thin lines represent autoregressive effects, statistical controls, or indicators of a latent variable. For Antisocial Behavior: A = Argue subscale, F = Fight subscale; For Need Satisfaction: A = Autonomy scale, C = Competence scale, R = Relatedness scale; For Need Frustration: A = Autonomy subscale, C = Competence subscale, R = Relatedness subscale; For Supportive Climate: R = Relatedness support subscale; I = improvement subscale; E = Effort subscale; For Conflictual Climate: C = Competition/Ability subscale; A = Ability subscale; COND = Experimental condition.

included them to provide a comprehensive test of the dual-process model.

5. Method

5.1. Participants

To be eligible for the study, a teacher needed to be a full-time PE teacher in Korea. Teachers were 49 full-time certified physical education (PE) teachers (31 males, 18 females) who taught in one of 49 different schools (20 middle, 29 high) dispersed throughout South Korea. All teachers were ethnic Korean. On average, teachers were 34.2 years old (*SD* = 4.3; *range* = 25–42) and had 7.0 years (*SD* = 4.2; *range* = 1–13) of PE teaching experience. All 49 teacher-participants completed all aspects of the study (i.e., the teacher retention rate was 100%), and each received the equivalent of $50 (unexpectedly) after the study in appreciation of their participation. Students were the 1487 ethnic Korean students in the classrooms of the 49 teachers (*M* class size = 30.4 students/class). On average, students were 15.4 years old (*SD* = 1.5, *range* = 13–18), including 789 (53.1%) females and 698 (46.9%) males, 608 (40.1%) middle and 879 (59.1%) high schoolers, and 752 (50.6%) in the experimental and 735 (49.4%) in the control condition. As to sample size adequacy, the estimation procedures needed within doubly latent multilevel structural equation models (DL-MLSEM) require a sample of 50 L2 units with at least 10 to 15 participants per unit (per classroom) (Lüdtke, Marsh, Robitzsch, & Trautwein, 2011; Morin, Blais, & Chenard-Poirier, 2021). The present sample generally met those requirements (49 teachers, 30.4 students/class), suggesting adequate statistical power.

5.2. Procedure

The University Research Ethics Committee of the first author's university approved the research protocol.

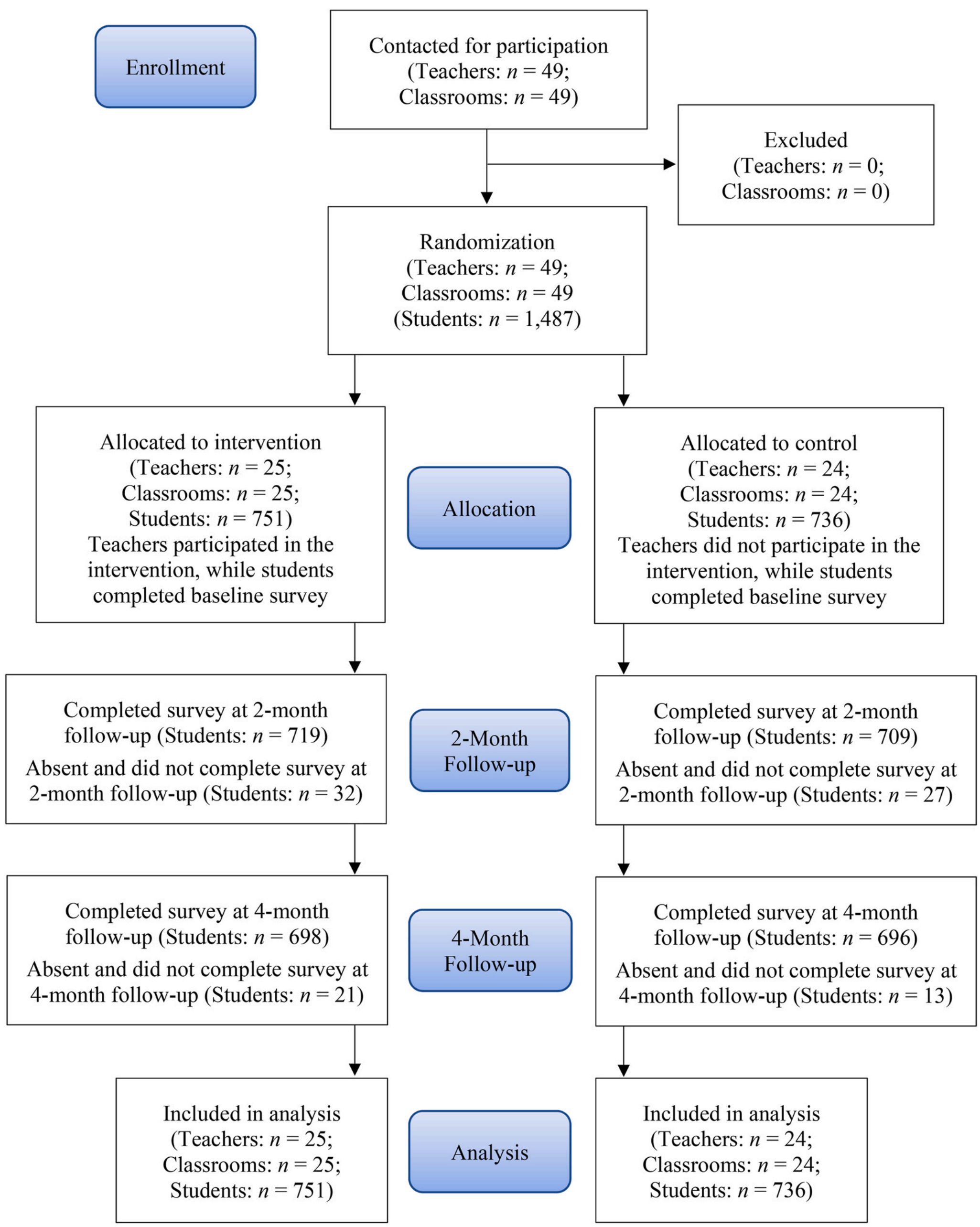

Figure 3. Intervention flowchart (CONSORT).

Research Design. The research design utilized a cluster randomized control trial with longitudinally-assessed dependent measures. Figure 3 provides a schematic overview and timeline of how PE teachers were recruited to participate, the computer-generated random assignment of teachers (and classrooms) to conditions, and the three waves of student-reported data collection. The teacher-focused intervention took place the week before classes began in the new academic year. The data collection took place in three waves in which students completed the same 4-page questionnaire at the beginning (T1; week 1), middle (T2; week 10), and end (T3; week 18) of the first semester of the 2017 academic year. Students completed the questionnaire in reference to that particular teacher and that particular class. The questionnaire began with a consent form, and students were assured that their responses would be confidential and used only for the purposes of the research study. At T1, students' questionnaire responses did not reflect an intervention effect (it was the first week of classes). Instead, student responses reflected their personal histories and expectations of what type of experiences are typical in a PE course (e.g., baseline need satisfaction and frustration, baseline classroom climate expectancies, baseline anticipated antisocial behavior).

During weeks 10 and 11, two trained raters from the research team visited each classroom to score teachers' in-class usage of autonomy-supportive and controlling instructional behaviors. The purpose of this rater-scored measure was to provide an objective test of the intervention effect (e.g., a manipulation check) on the quality of teachers' actual classroom instruction in terms of how autonomy-supportive and how controlling it was. To make these ratings, the two raters came to the class unannounced 5–10 min before its start, did not know into which group the observed teacher had been randomly assigned, and made independent ratings. However, because of the wide geographical diversity of the 49 participating schools and the need to collect the teacher ratings during the same time period (week 10 or 11), raters could only visit and make behavioral ratings of 38 of the 49 (78%) participating teachers.

Autonomy-Supportive Teaching Intervention. We provided teachers in the experimental condition with a 3-part, 8-h autonomy-supportive teaching intervention. The contents, activities, and step-by-step procedures of the autonomy-supportive teaching intervention followed previously published 3-part ASIPs (Cheon & Reeve, 2015; Cheon, Reeve, Lee, et al., 2019; Cheon et al., 2016, 2018, 2019)—with one new addition to Part 1, as described below.

Part 1 was a 3-h, morning presentation one week before the school year began. It began with a reflective warm-up activity to help teachers become aware of their own autonomy-supportive and controlling teaching tendencies. It then featured a media-rich Power-Point presentation on autonomy-supportive teaching, empirical evidence on the benefits of autonomy support, and PE-specific examples of the seven recommended autonomy-supportive instructional behaviors—namely, take the students' perspective, invite students to pursue their personal interests, present learning activities in need-satisfying ways, provide explanatory rationales, acknowledge negative feelings, rely on invitational language, and display patience (Reeve & Cheon, 2021). The new addition was that, during the last 20 min of Part 1, teachers observed examples and models of how each recommended instructional behavior could also be delivered at the teacher-to-whole-class level. For instance, teachers were shown not only how to take the perspective of and provide an explanatory rationale to an individual student but also how to take the perspective of and provide an explanatory rationale to the class as a whole (see Kaplan & Assor, 2012).

Part 2 was a same-day 3-h, afternoon workshop. It focused on the practical "how to" of the seven recommended instructional behaviors. Each autonomy-supportive act of instruction was first described and modeled (via a series of brief, professionally-produced video clips) and then practiced, refined, and discussed until teachers felt sufficiently skilled to try out that act of instruction in their own classrooms with their own students. In the final half-hour, teachers practiced replacing their existing controlling instructional behaviors with alternative/replacement autonomy-supportive behaviors (replacing pressuring language with invitational language). This aspect of the workshop was designed to help teachers become not only more autonomy-supportive but also less controlling during their instruction.

Part 3 took place one month into the semester (after teachers had actual classroom experience with autonomy-supportive teaching). It featured a peer-to-peer group discussion about teachers' early-semester experiences with autonomy-supportive teaching. Teachers both gave and received peer-informed instructional help, tips, and strategies on how to become more autonomy-supportive and less controlling toward students—especially in terms of how to enact autonomy-supportive teaching in those teaching situations that were of special concern to individual teachers (e.g., students sit passively on the sidelines, a student plays games on her smartphone).

5.3. *Classroom observers' rating sheets*

Raters used the Behavior Rating Scale (BRS; Cheon et al., 2018) to make their mid-semester ratings of teachers' in-class usage of autonomy-supportive and controlling instructional behaviors, The BRS features a 1–7 unipolar scale (1 = *not at all*, 7 = *very much*) to ask raters to score individual instructional behaviors (e.g., takes the students' perspective, uses pressuring language). These scores are then averaged into (1) an overall "rater-scored autonomy-supportive teaching" score ($\alpha = 0.94$) and (2) an overall "rater-scored controlling teaching" score ($\alpha = 0.90$).

5.4. *Students' questionnaire measures*

The student-reported measures all used the same 1–7 bipolar scale (1 = *Strongly Disagree*, 7 = *Strongly Agree*). We assessed students' psychological needs with questionnaire items that used the individual student as the referent ("I feel that I do PE activities because I want to."), because we conceptualized need satisfaction and need frustration as individually experienced phenomena. We assessed students' perceptions of their teachers' motivating style with questionnaire items that used the teacher as the referent (e.g., "My PE teacher encourages me to ask questions."), because we conceptualized teacher behavior as a group experienced phenomenon. We assessed students' perceptions of the classroom climates and antisocial behavior with questionnaire items that used the class as the referent (e.g., "During this PE class, most students fight."), because we conceptualized these as group experienced phenomena. For each measure, we had available a previously back-translated and successfully validated Korean version of the original English-language questionnaire (e.g., Cheon & Song, 2011; Jang, Kim, & Reeve, 2016; Song et al., 2016).

We calculated the inter-item (α) and inter-rater (*ICC1*, *ICC2*) reliability statistics for each measure across all three waves of data. The alpha coefficient (α) reports the internal consistency of the individual items on a questionnaire, and values above 0.70 reflect reasonable internal consistency. The Inter-class Correlation Coefficient *ICC1* statistic reports the proportion of the variance in the dependent measure attributable to class membership (i.e., the shared agreement among students in that class on the level of that dependent measure). An *ICC1* value above 0.100 indicates a "shared perception" (group consensus) that can be studied in a multilevel analysis as an L2 phenomenon (Lüdtke et al., 2008). These L2 variables have a different meaning and a different interpretation than do the L1 variables (i.e., a "group reality" vs. an "individual perception" to use Morin et al.'s, 2021, terminology), and our hypotheses were exclusive to these L2 classroom phenomena and processes. The *ICC2* statistic represents the reliability of the aggregated *ICC1* score among the 30 or so students in the class. An *ICC2* value above 0.700 suggests a reasonably reliable aggregate group (class) score (Klein & Kozlowski, 2000). These ICC statistics for the teacher-referenced and classroom-referenced measures can be relatively low at T1, because students have little shared history at week 1.

However, the *ICC* statistics were expected to rise above 0.100 at T2 and T3 as students spent more time together.

Perceived Autonomy-Supportive and Controlling Teaching. We assessed students' perceived autonomy-supportive teaching with the 6-item Learning Climate Questionnaire (LCQ; Black & Deci, 2000). Students' reports on the LCQ (e.g., "My PE teacher listens to how I would like to do things.") were internally consistent across the three waves of data collection (αs at T1, T2, and T3 were 0.89, 0.91, and 0.93, respectively), showed a rising within-class consensus (*ICC1s* = 0.058, 0.167, and 0.144), and a reasonably high reliability of that consensus (*ICC2s* = 0.654, 0.882, and 0.836). We assessed students' perceived controlling teaching with the 4-item Controlling Teacher Questionnaire (CTQ; Jang et al., 2009). Students reports on the CTQ (e.g., "My PE teacher tries to control everything I do.") were internally consistent (αs = 0.80, 0.85, and 0.85), showed a rising within-class consensus (*ICC1s* = 0.058, 0.151, and 0.121), and a reasonably high reliability of that consensus (*ICC2s* = 0.667, 0.844, and 0.808).

Need Satisfaction and Need Frustration. We assessed autonomy, competence, and relatedness satisfaction with three separate scales. For autonomy satisfaction, we used the 5-item Perceived Autonomy scale (e.g., "In this PE class, I can decide which activities I want to do"; Standage et al., 2006); students' reports were internally consistent (αs = 0.85, 0.90, and 0.92), showed a modest within-class consensus (*ICC1s* = 0.057, 0.104, and 0.110), and a modest reliability of that consensus (*ICC2s* = 0.648, 0.779, and 0.791). For competence satisfaction, we used the 4-item Perceived Competence scale from the Intrinsic Motivation Inventory (e.g., "I think I am pretty good at PE activities"; McAuley et al., 1989); students' reports were internally consistent (αs = 0.88, 0.90, and 0.91), showed a reasonable within-class consensus (*ICC1s* = 0.104, 0.120, and 0.128), and a reasonable reliability of that consensus (*ICC2s* = 0.779, 0.805, and 0.817). For relatedness satisfaction, we used the 4-item Sense of Relatedness scale (e.g., "When I am with my PE classmates, I feel accepted"; Furrer & Skinner, 2003); students' reports were internally consistent (αs = 0.82, 0.88, and 0.88), showed a low within-class consensus (*ICC1s* = 0.044, 0.070, and 0.068), and only a modest reliability of that consensus (*ICC2s* = 0.582, 0.695, and 0.691).

We assessed autonomy, competence, and relatedness frustration with the 12-item Psychological Need Thwarting Scale (PNTS; Bartholomew et al., 2011b), which includes three 4-item subscales to assess autonomy, competence, and relatedness frustration. For autonomy frustration (e.g., "In this PE class, I feel pushed to behave in certain ways"), students' reports were internally consistent (αs = 0.78, 0.86, and 0.88), showed a low within-class consensus (*ICC1s* = 0.029, 0.089, and 0.061), and only a modest reliability of that consensus (*ICC2s* = 0.478, 0.748, and 0.664). For competence frustration (e.g., "In this PE class, there are situations where I am made to feel inadequate"), students' reports were internally consistent (αs = 0.87, 0.92, and 0.93), showed a low within-class consensus (*ICC1s* = 0.021, 0.075, and 0.050), and only a modest reliability of that consensus (*ICC2s* = 0.391, 0.713, and 0.615). For relatedness frustration (e.g., "In this PE class, I feel I am rejected by those around me"), students' reports were internally consistent (αs = 0.88, 0.94, and 0.94), showed a low within-class consensus (*ICC1s* = 0.062, 0.124, and 0.042), and only a modest reliability of that consensus (*ICC2s* = 0.667, 0.812, and 0.569).

Supportive and Conflictual Classroom Climates. We used the 21-item, 5 subscale Peer Motivational Climate in Youth Sport Questionnaire (PeerMCYSQ; Ntoumanis & Vazou, 2005) to assess the two classroom climates. We assessed the supportive climate with the 3-item Relatedness Support scale, the 4-item Improvement scale, and the 5-item Effort scale. In contrast, we assessed the conflictual climate with the 4-item Intrateam Conflict scale and the 5-item Competition/Ability scale. For relatedness support (e.g., "During this PE class, most students care about everyone's opinions."), students' reports were internally consistent (αs = 0.87, 0.93, and 0.94), showed a rising within-class consensus (*ICC1s* = 0.083, 0.123, and 0.138), and an acceptable reliability of that consensus (*ICC2s* = 0.734, 0.809, and 0.829). For improvement (e.g., "During this PE class, most students teach their classmates new things."), students' reports were internally consistent (αs = 0.90, 0.95, and 0.96), showed a rising within-class consensus (*ICC1s* = 0.102, 0.171, and 0.137), and an acceptable reliability of that consensus (*ICC2s* = 0.775, 0.862, and 0.828). For effort (e.g., "During this PE class, most students encourage their classmates to try their hardest."), students' reports were internally consistent (αs = 0.88, 0.93, and 0.94), showed a rising within-class consensus (*ICC1s* = 0.088, 0.160, and 0.149), and an acceptable reliability of that consensus (*ICC2s* = 0.745, 0.853, and 0.842). For intrateam conflict (e.g., "During this PE class, most students criticize their classmates when they make mistakes."), students' reports were internally consistent (αs = 0.89, 0.94, and 0.95), showed a rising within-class consensus (*ICC1s* = 0.107, 0.144, and 0.133), and an acceptable reliability of that consensus (*ICC2s* = 0.785, 0.836, and 0.823). For competition/ability (e.g., "During this PE class, most students looked pleased when they do better than their classmates."), students' reports were internally consistent (αs = 0.79, 0.85, and 0.86), showed a moderate within-class consensus (*ICC1s* = 0.069, 0.099, and 0.107), and an acceptable reliability of that consensus (*ICC2s* = 0.692, 0.769, and 0.785).

Antisocial Behavior. We used the Prosocial and Antisocial Behavior in PE scale (PABPE; Cheon et al., 2017) to assess antisocial behavior, because it was developed specifically for the PE classroom setting and used "my classmates" as the sentence stem referent. We assessed antisocial behavior with both the 2-item Fight scale and the 2-item Argue scale. For fight (e.g., "In this PE class, my classmates fight."), students' reports were internally consistent (αs = 0.68, 0.78, and 0.80), showed a reasonable within-class consensus (*ICC1s* = 0.073, 0.120, and 0.078), and an acceptable reliability of that consensus (*ICC2s* = 0.705, 0.805, and 0.720). For argue (e.g., "In this PE class, my classmates argue."), students' reports were internally consistent (αs = 0.80, 0.85, and 0.86), showed a moderate within-class consensus (*ICC1s* = 0.045, 0.101, and 0.070), and a moderate reliability of that consensus (*ICC2s* = 0.591, 0.773, and 0.695).

5.5. *Data analyses*

Preliminary analyses showed that values for skewness and kurtosis for all the rater-scored and student-reported dependent measures (and their subscales) at T1, T2, and T3 were all less than |2|, indicating little deviation from normality.

Rater-scored Instructional Behaviors. To test for the effect of experimental condition on rater-scored autonomy-supportive and controlling instructional behaviors (i.e., the intervention effect), the unit-of-analysis was the teacher ($N = 49$) and the statistical test was a 2-group independent *t*-test. To provide effect size information, we used Cohen's *d* (Cohen, 1988).

Doubly Latent Multilevel Structural Equation (DL-ML-SEM) Analysis. All student-reported data had a 3-level hierarchical structure with repeated measures (3 waves, $N = 4461$) nested within students ($N = 1487$), nested within teachers/classrooms ($k = 49$). To analyze these data, we used a DL-ML-SEM analysis (Marsh et al., 2011; Morin et al., 2014, 2021). In doing so, we used Mplus 8.3 (Muthen & Muthen, 2019) with the maximum likelihood-robust estimator (MLR) and full information maximum likelihood (FIML) estimation procedures for handling missing data. To evaluate model fit, we used the following goodness-of-fit statistics: Root-mean-square error of approximation (RMSEA), standardized root-mean-square residual (SRMR), comparative fit index (CFI), and Tucker-Lewis index (TLI). For RMSEA and SRMR, adequate and excellent fit are reflected by values lower than 0.08 and 0.06; for CFI and TLI, adequate and excellent fit are reflected by values greater than 0.90 and 0.95 (Marsh, Hau, & Grayson, 2005).

In a DL-ML-SEM analysis, students' questionnaire responses are used to create latent variables at both the student (L1) and classroom (L2) levels. Effects at the L1 and L2 levels can be studied as distinct effects, because a DL-ML-SEM analysis disaggregates the L1 and L2 components of students' ratings to control for unreliability in the aggregation of the L2 ratings and to control for the sampling error from the 30 or so different students in each class. Different interpretations of the L1 and L2 values are applied to questionnaire items with an individual referent (e.g., "I feel that I do PE activities because I want to.") vs. a classroom referent (e.g., "During this PE class, most students make their classmates feel valued."), as explained next.

For constructs assessed with an individual referent (i.e., need satisfaction, need frustration), student ratings are used to form traditional latent variables at the L1 level that show student-to-student variation (e.g., I feel high need satisfaction). These ratings are also aggregated at the L2 level to provide additional meaning as a "context" variable (Marsh et al., 2011). For constructs assessed with a classroom referent (i.e., perceived autonomy-supportive and controlling teaching, supportive and conflictual classroom climates, and antisocial behavior), student ratings are aggregated at the L2 level to extract the "shared agreement" among the 30 or so students in the class. These L2 latent variables provide a "climate" interpretation (Marsh et al., 2011). That is, students' responses are parsed into two separate components—one to represent the group reality (the aggregated L2 variable) and one to represent within-class student-to-student differences in their perception of that group reality (the residual L1 variable). The L2 component has a clear meaning (a gauge of the prevailing classroom climate), while the L1 component reflects inter-individual differences in perceptions of the L2 group reality. These disaggregated L2 and L1 variables can correlate weakly, can predict different outcomes, and L1 effects do not necessarily generalize to parallel L2 effects (Marsh et al., 2011, 2009; Morin et al., 2021; Richey, Bernacki, Belenky, & Nokes-Malach, 2018).

In a DL-ML-SEM analysis, it is important (for interpretative considerations) to establish multilevel measurement isomorphism (Morin et al., 2021). Isomorphism means metric invariance, or the equality of the factor loadings across the L1 and L2 levels. To test for isomorphism, the factor loadings of the indicators (of the latent variables) are all fixed to their L1 and T1 values. If the measurement model that constrain these indicators to be invariant across both level and time shows little or no decrement in fit (according to the goodness-of-fit statistics) compared to the measurement model in which the indicators are free to vary, then measurement isomorphism is verified (Marsh et al., 2011).

Tests of the Student-reported Intervention Effect. Students reported their perceptions of how autonomy-supportive and how controlling they perceived their teacher to be (using the LCQ and CTQ, respectively). In both analyses, we conducted a growth model (within the DL-ML-SEM framework) that regressed the latent variable dependent measure on the slope of the T1, T2, and T3 scores (weighted as 0, 1, 2). Experimental condition (control = 0, experimental = 1) was the critical independent variable, gender (male = 0, female = 1) was an L1 covariate, and grade level (middle school = 0, high school = 1) and class size were L2 covariates. For these analyses, we were simply interested in evaluating for a significant effect of experimental condition on the linear growth (longitudinal change) in each L2 dependent measure—essentially a condition × time interaction effect.

Test of the Measurement Models. We tested the fit of two measurement models—one for the Needs Only Model and a second for the Needs and Climates Model. The Needs Only measurement model included 16 indicators to create 6 L1 latent variables and 3 L2 latent variables (see Figure 1). The Needs and Climates measurement model included 26 indicators to create 10 L1 latent variables and 5 L2 latent variables (see Figure 2). In both cases, we first tested a measurement model that allowed parameter estimates to vary freely, while we second tested a measurement model that included restraints to test for multilevel and multiwave measurement isomorphism.

Test of the Hypothesized Models. We tested the fit of two hypothesized models—one for the Needs Only Model and a second for the Needs and Climates Model. To create these models, we supplemented the corresponding measurement models with experimental condition as an uncentered predictor (control condition = 0, experimental condition = 1), gender as a group mean-centered L1 covariate (0 = male, 1 = female), and grade level (0 = middle school, 1 = high school) and class size (*range* = 22 to 37) as two grand mean-centered L2 covariates. The Mplus syntax for the test of the Needs and Climates Model appears in the Supplemental Materials.

Mediation Analyses. The hypothesized models shown in Figures 1 and 2 are both mediation models, so we tested for mediation effects. The typical procedure to test for such mediation effects is to use resampling methods to generate bias-corrected confidence intervals, but this conventional bootstrapping method cannot be applied to multilevel modeling, because the assumption of independence of observations is violated when using nested or clustered data (Preacher & Selig, 2012). Accordingly, we utilized a Monte Carlo approach to resampling that allowed us to construct the appropriate confidence intervals. To do so, we used Selig and Preacher's (2008) web-based utility[2] to generate and run R code for simulating the sampling distribution of each indirect effect (20,000 values). If the 95% CI from this simulation excludes zero, then the indirect effect test is significant ($p < .05$).

6. Results

6.1. Intervention effect [3]

Figure 4 shows the effect that teacher participation in the intervention (i.e., experimental condition) had on rater-scored autonomy-supportive and controlling instructional behaviors and students' perceived autonomy-supportive and controlling teaching.

Rater-Scored Autonomy-Supportive and Controlling Teaching. As shown in Figure 4a, at mid-semester, raters scored teachers in the experimental condition (M = 5.47, SD = 0.29) as engaging in more autonomy-supportive instructional behaviors than they scored teachers in the control condition (M = 4.27, SD = 0.47), $t(36) = 9.43$, $p < .001$, d = 3.03. Similarly, raters scored teachers in the experimental condition (M = 2.47, SD = 0.30) as engaging in fewer controlling instructional behaviors than they scored teachers in the control condition (M = 3.36, SD = 0.46), $t(36) = 7.14$, $p < .001$, $d = 2.29$.

Students' Perceived Autonomy-Supportive Teaching. The multilevel (DL-ML-SEM) linear growth model to explain longitudinal gains in students' L2 perceived autonomy-supportive teaching as a function of experimental condition fit the data reasonably well, $X^2(308) = 1332.33$, $p < .001$, $RMSEA = 0.047$, $SRMR = 0.033$, $CFI = 0.947$, $TLI = 0.935$. As illustrated in Figure 4b, over the course of the semester, experimental condition predicted the growth in how autonomy-supportive students perceived their teachers to be ($\beta = 0.85$, $SE\ B = 0.07$, $t = 12.94$, $p < .001$; $R^2 = 0.73$), as students of teachers in the experimental group reported a greater T3 growth in perceived autonomy-supportive teaching than did students of teachers in the control group ($\Delta\ Ms = +1.00$ vs. $+0.11$).

Perceived Controlling Classroom Climate. The multilevel linear growth model to explain longitudinal declines in students' L2 perceived controlling teaching as a function of experimental condition fit the data reasonably well, $X^2(128) = 402.79$, $p < .001$, $RMSEA = 0.038$, $SRMR = 0.027$, $CFI = 0.963$, $TLI = 0.947$. As illustrated in Figure 4c,

[2] http://quantpsy.org/medmc/medmc.htm.

[3] We used observers' ratings and students' perceptions of teachers' motivating styles as two independent tests of the experimental manipulation. We checked to see whether observers' and students' middle-of-semester (T2) ratings corresponded with each other, and they did. Observers' ratings of teachers' autonomy-supportive instructional behavior scored at the teacher level ($k = 38$, $M = 4.90$, $SD = 0.72$) significantly predicted (i.e., agreed with) T2 perceived autonomy-supportive teaching scored at the student level (N = 1,131, M = 5.24, $SD = 1.08$): $B = 0.42$, $SE = 0.09$, $t(36) = 4.51$, $p < .001$. Similarly, observers' ratings of teachers' controlling instructional behavior scored at the teacher level ($k = 38$, $M = 2.89$, $SD = 0.59$) significantly predicted T2 perceived controlling teaching scored at the student level (N = 1,131, M = 2.54, SD = 1.19): $B = 0.39$, $SE = 0.09$, $t(36) = 3.40$, $p = .002$.

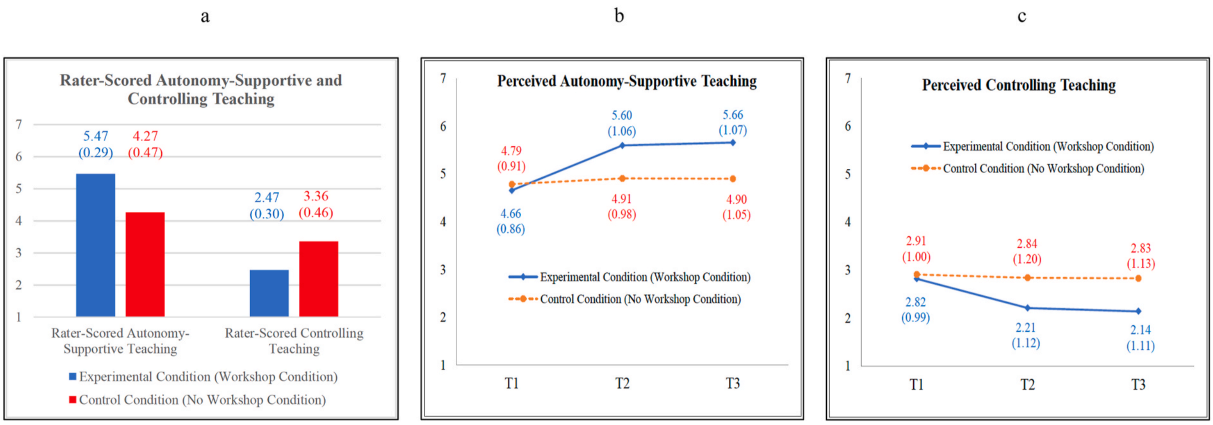

Figure 4. *Means and Standard Deviations (in Parentheses) for the Three Manipulation Checks: Rater-Scored Autonomy-Supportive and Controlling Teaching (left panel); Students' Perceived Autonomy-Supportive Teaching (center panel), and Students' Perceived Controlling Teaching (right panel) Broken Down by Experimental Condition and Time of Assessment.* T1 = Time 1, T2 = Time 2, T3 = Time 3.

experimental condition predicted the decline in how controlling students perceived their teachers to be ($\beta = -0.75$, $SE\ B = 0.14$, $t = 5.58$, $p < .001$; $R^2 = 0.69$), as students of teachers in the experimental group reported a greater T3 decline in perceived controlling teaching than did students of teachers in the control group ($\Delta\ Ms = -0.68$ vs. -0.08).

6.2. Test of the needs only model

We first tested the Needs Only measurement model. This measurement model fit the data only adequately, $X^2(190) = 1850.40$, $p < .001$, $RMSEA = 0.077$, $SRMR = 0.064$, $CFI = 0.903$, and $TLI = 0.877$, though factor loadings for the multi-indicator latent constructs were all substantial and statistically significant ($p < .001$). We next constrained the indicators to be invariant across both level and time (by fixing the factor loadings for each latent construct to their L1 and T1 values), and this invariant measurement model also fit the data only adequately, $X^2(200) = 1827.13$, $p < .001$, $RMSEA = 0.074$, $SRMR = 0.065$, $CFI = 0.905$, $TLI = 0.886$, though it showed no decrement in the fit indices.

We next tested the structural model by adding experimental condition, the hypothesized paths (see Fig. 1), and the statistical controls. This structural model fit the data adequately, $X^2(257) = 2044.08$, $p < .001$, $RMSEA = 0.068$, $SRMR = 0.072$, $CFI = 0.899$, and $TLI = 0.881$. In the prediction of L2 T3 antisocial behavior (overall $R^2 = 0.51$), L2 T2 need frustration (H3) was not an individually significant predictor ($B = 0.32$, $SE\ B = 0.29$, $\beta = 0.36$, $t = 1.09$, $p = .278$), after controlling for experimental condition ($\beta = -0.39$, $p = .068$), grade level ($\beta = 0.09$, $p = .527$), and class size ($\beta = -0.15$, $p = .341$). In the prediction of L1 T3 antisocial behavior (overall $R^2 = 0.24$), L1 T2 need frustration was an individually significant predictor ($B = 0.30$, $SE\ B = 0.04$, $\beta = 0.30$, $t = 7.45$, $p < .001$), controlling for L1 T1 need frustration ($\beta = 0.08$, $p = .053$), L1 T1 antisocial behavior ($\beta = 0.26$, $p < .001$), and gender ($\beta = -0.06$, $p = .028$).

We did not conduct a follow-up mediation analysis of the Needs Only Model, because L2 T2 need frustration was not an individually significant predictor of L2 T3 antisocial behavior. In a supplemental analysis, we conducted a cross-over analysis by adding L2 (T2) and L1 (T1, T2) need satisfaction as predictors. This cross-over structural model fit the data about the same, $X^2(256) = 2041.87$, $p < .001$, $RMSEA = 0.068$, $SRMR = 0.072$, $CFI = 0.899$, and $TLI = 0.880$, but not any better than the original needs only model, $\Delta X^2 (\Delta df = 1) = 2.21$, $p = .137$. At the L2 level, T2 need satisfaction did not individually predict L2 T3 antisocial behavior ($\beta = -0.04$, $p = .902$), L2 T2 need frustration remained a non-significant predictor ($\beta = 0.37$, $p = .280$), and the variance explained in L2 T3 antisocial behavior rose only from 0.51 to 0.52.

6.3. Test of the needs and climates model

The measurement model underlying the Needs and Climates Model fit the data reasonably well, $X^2(543) = 2849.27$, $p < .001$, $RMSEA = 0.053$, $SRMR = 0.047$, $CFI = 0.922$, and $TLI = 0.907$. Factor loadings for indicators of the latent constructs were all substantial and statistically significant ($p < .001$). After constraining the indicators to be invariant across both level and time, the invariant measurement model continued to fit the data well and showed little or no decrement in the fit indices, $X^2(559) = 2864.27$, $p < .001$, $RMSEA = 0.053$, $SRMR = 0.048$, $CFI = 0.922$, $TLI = 0.910$, thereby establishing measurement isomorphism across both levels (L1 and L2) and waves (T1, T2, and T3).

We next tested the hypothesized model, and it fit the data reasonably well, $X^2(661) = 3206.99$, $p < .001$, $RMSEA = 0.051$, $SRMR = 0.061$, $CFI = 0.917$, and $TLI = 0.905$. The descriptive statistics and intercorrelations among experimental condition, the L1 and L2 latent variables, and the L1 and L2 covariates appear in Table 1. The standardized beta weights for the structural paths, autoregressive effects, and statistical controls appear in Figure 5.

Experimental condition significantly predicted all four L2 T2 dependent measures: T2 need satisfaction (H1: $B = 0.54$, $SE\ B = 0.06$, $\beta = 0.84$, $t = 8.34$, $p < .001$); T2 need frustration (H2: $B = -0.40$, $SE\ B = 0.08$, $\beta = -0.66$, $t = 5.28$, $p < .001$); T2 supportive climate (H4: $B = 0.59$, $SE\ B = 0.08$, $\beta = 0.77$, $t = 7.31$, $p < .001$); and T2 conflictual climate (H5: $B = -0.47$, $SE\ B = 0.09$, $\beta = -0.68$, $t = 5.33$, $p < .001$), after controlling for grade level and class size.

In the prediction of L2 T3 antisocial behavior (overall $R^2 = 0.86$), L2 T2 conflictual climate (H6) was an individually significant predictor ($B = 0.78$, $SE\ B = 0.24$, $\beta = 1.01$, $t = 3.23$, $p < .001$),[4] while L2 T2 need frustration (H3) was not ($B = -0.20$, $SE\ B = 0.20$, $\beta = -0.23$, $t = 0.98$, $p = .325$), after controlling for experimental condition ($\beta = -0.09$, $p = .569$), grade level ($\beta = 0.16$, $p = .162$), and class size ($\beta = -0.09$, $p = .481$). In the prediction of L1 T3 antisocial behavior (overall $R^2 = 0.26$), L1 T2 peer conflict ($B = 0.28$, $SE\ B = 0.06$, $\beta = 0.28$, $t = 4.86$, $p < .001$), L1 T2 need frustration ($B = 0.18$, $SE\ B = 0.03$, $\beta = 0.17$, $t = 5.33$, $p < .001$), and L1 T1 need frustration ($B = 0.10$, $SE\ B = 0.04$, $\beta = 0.10$, $t = 2.64$, $p = .008$), but not L1 T2 peer conflict ($B = -0.07$, $SE\ B = 0.05$, $\beta =$

[4] The $\beta = 1.01$ value for L2 T2 conflictual climate to predict L2 T3 antisocial behavior represents a slight suppressor effect. L2 T2 conflictual climate and L2 T2 need frustration were two positively correlated predictors ($r = 0.77$, see Table 1). When the $\beta = -0.23$ for L2 T2 need frustration to L2 T3 antisocial behavior (see Fig. 5) flipped from a positive relation ($r = 0.63$, see Table 1), the L2 T3 antisocial behavior then artificially rose from and underlying $r = 0.91$ (see Table 1) to a $\beta = 1.01$ (see Fig. 5) that represented a suppressor effect.

Table 1
L1 and L2 descriptive statistics and intercorrelation matrices for experimental condition, the latent variables, and the statistical controls included in the test of the hypothesized model.

Student (L1) Level												
Variable	*M*	*SD*	2.	3.	4.	5.	6.	7	8.	9.	10.	11.
1. T1 Need Satisfaction	4.85	0.82	-.65	.73	-.40	-.40	.56	-.22	.34	-.19	-.23	n/a
2. T1 Need Frustration	2.45	0.89	–	-.49	.49	.60	-.37	.34	-.23	.23	.33	n/a
3. T1 Supportive Climate	4.81	0.82		–	-.54	-.46	.41	-.17	.47	-.26	-.26	n/a
4. T1 Conflictual Climate	3.45	0.98			–	.59	-.22	.17	-.25	.47	.28	n/a
5. T1 Antisocial Behavior	2.84	1.07				–	-.22	.21	-.22	.28	.34	n/a
6. T2 Need Satisfaction	5.28	0.94					–	-.51	.74	-.39	-.36	-.05
7. T2 Need Frustration	2.20	1.13						–	-.38	.48	.38	-.06
8. T2 Supportive Climate	5.29	1.06							–	-.50	-.39	-.03
9. T2 Conflictual Climate	3.28	1.28								–	.42	-.06
10. T3 Antisocial Behavior	2.69	1.33									–	-.08
11. Student Gender	0.53	0.50										–

Classroom (L2) Level									
Variable	*M*	*SD*	2.	3.	4.	5.	6.	7.	8.
1. Experimental Condition	0.51	0.50	.81	-.65	.76	-.67	-.61	.02	-.09
2. T2 Need Satisfaction	5.28	0.38	–	-.63	.96	-.72	-.61	-.12	.08
3. T2 Need Frustration	2.20	0.42		–	-.51	.77	.63	-.04	-.05
4. T2 Supportive Climate	5.30	0.49			–	-.76	-.66	-.07	.04
5. T2 Conflictual Climate	3.28	0.54				–	.91	-.04	-.07
6. T3 Antisocial Behavior	2.69	0.45					–	.11	-.14
7. Grade Level	0.59	0.49						–	.00
8. Class Size	30.4	0.6							–

Numbers in the upper half of the table display correlations at the student level (L1; $N = 1487$ students); any L1 correlation $r \geq 0.06$, $p < .05$. Numbers in the lower half of the table display correlations at the classroom level (L2; $k = 49$ classrooms); any L2 correlation $r \geq 0.29$, $p < .05$. n/a = not applicable in a multilevel analysis.

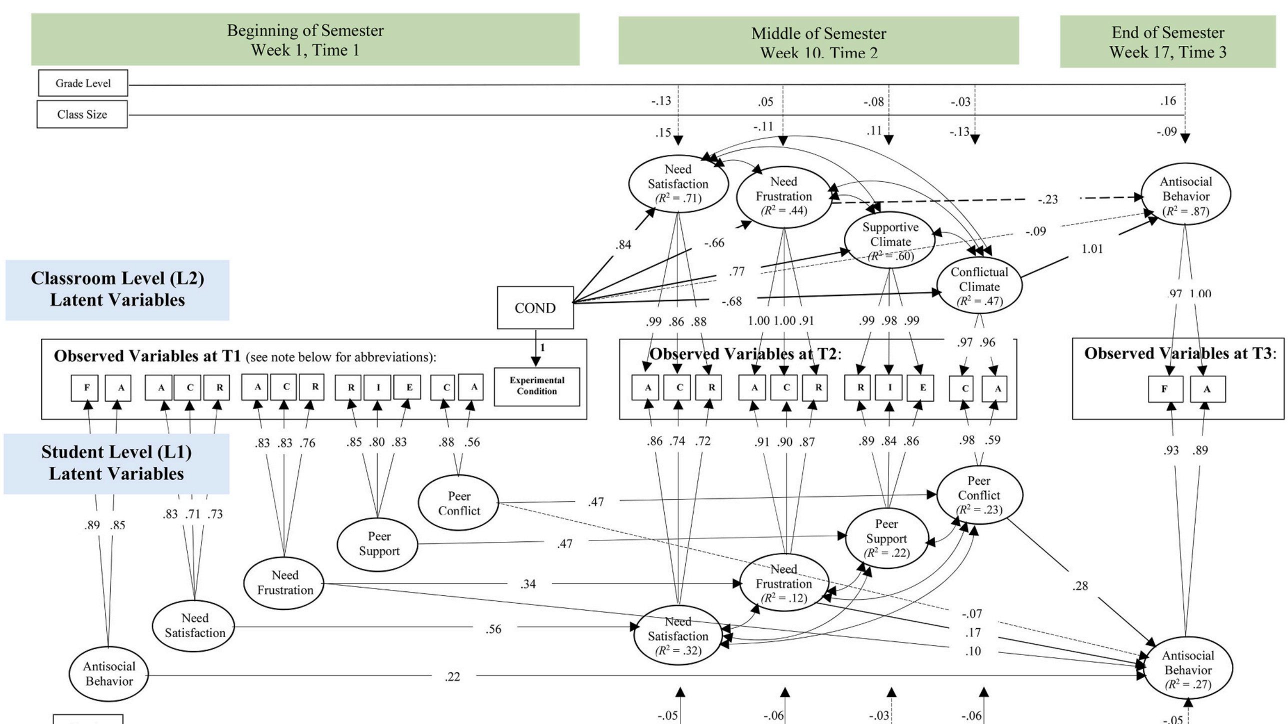

Figure 5. Results of the Test of the Needs and Climates Hypothesized Model to Predict T3 Antisocial Behavior
Note. Solid lines represent significant paths, $p < .05$; dashed lines represent non-significant paths. Numbers that overlay the solid and dashed lines are standardized beta weights. For Antisocial Behavior: A = Argue subscale, F = Fight subscale; For Need Satisfaction: A = Autonomy scale, C = Competence scale, R = Relatedness scale; For Need Frustration: A = Autonomy subscale, C = Competence subscale, R = Relatedness subscale; For Supportive Climate: R = Relatedness support subscale; I = improvement subscale; E = Effort subscale; For Conflictual Climate: C = Competition/Ability subscale; A = Ability subscale; COND = Experimental condition.

-0.07, $t = 1.57$, $p = .116$), were individually significant predictors, after controlling for L1 T1 antisocial behavior ($\beta = 0.22$, $p < .001$) and gender ($\beta = -0.05$, $p = .073$).

In the mediation analysis, the confidence interval for the indirect effect of experimental condition on L2 T3 antisocial behavior via L2 T2 conflictual climate did not include zero (0.044, 0.326), thereby confirming T2 conflictual climate as a mediator. The confidence interval for the indirect effect of experimental condition on L2 T3 antisocial behavior via L2 T2 need frustration did include zero (−0.109, 0.035), thereby disconfirming T2 need frustration as a mediator.

Lastly, we conducted a cross-over analysis by adding the L2 (T2) and L1 (T1, T2) need satisfaction and supportive climate predictors. This cross-over structural model fit the data reasonably well, $X^2(659)$ = 3204.20, $p < .001$, $RMSEA = 0.051$, $SRMR = 0.060$, $CFI = 0.917$, and $TLI = 0.905$, but no better than the hypothesized model, ΔX^2 ($\Delta df = 2$) = 2.79, $p = .248$. At the L2 level, neither T2 need satisfaction nor T2 supportive classroom climate were able to individually predict L2 T3 antisocial behavior, while L2 T2 conflictual climate remained a significant predictor, L2 T2 need frustration remained a non-significant predictor, and the variance explained in L2 T3 antisocial behavior rose only from 0.86 to 0.87.

7. Discussion

Autonomy-supportive teaching interventions produce student benefits, and they do so because they enable greater need satisfaction and lesser need frustration. The evidence supporting this claim and its interpretation are strong (Reeve & Cheon, 2021), but we initiated the present investigation on the premise that this explanatory model overlooks the key facilitating effect of an improved classroom climate. The purpose of the present study, therefore, was to test the extent to which a "Needs and Climates" model would better explain reduced antisocial behavior than a "Needs Only" model. The key finding was that the intervention reduced antisocial behavior largely because it diminished a conflictual classroom climate.

7.1. Autonomy-supportive teaching improves the classroom climate

Participation in the autonomy-supportive teaching intervention gave PE teachers the skill they needed to foster a more supportive and less conflictual classroom climate. These effects were so large (in terms of magnitude; see β s in Fig. 5) that they raise the question of how past researchers overlooked them. Researchers overlooked these effects because they relied exclusively on a student unit-of-analysis. Such investigations measure and put the spotlight on the individual student experience. However, investigations that rely on a classroom unit-of-analysis measure and put the spotlight on the environment that surrounds those individual students (Marsh et al., 2011; Smith, 2003).

As shown in the upper center part of Figure 5, autonomy-supportive teachers generated a PE classroom environment in which students largely agreed that they were surrounded by (1) classmates who collectively experienced high need satisfaction (H1), (2) classmates who collectively experienced low need frustration (H2), (3) a classroom climate rich in supportive peer-to-peer interactions and relationships (H4), and (4) a classroom climate lacking conflictual peer-to-peer interactions and relationships (H5). The first two effects show that autonomy-supportive teachers facilitated favorable needs-based processes. In comparison, the latter two effects show that autonomy-supportive teachers facilitated favorable group- and relationship-based processes.

Class-wide norms are largely created and maintained by student-to-student interactions and relationships. Teachers in the control condition had little effect on these peer-to-peer group dynamics. They did little to bend the classroom norms and peer interactions either toward interpersonal support or away from interpersonal conflict. However, teachers in the experimental condition provided instruction in a way that did alter these classroom norms and peer interactions toward both greater support and lesser conflict. Autonomy-supportive teachers did this during their whole-class instruction by taking their students' perspective, providing explanatory rationales for each teacher request, acknowledging and accepting students' expressions of negative feelings, relying on invitational language, displaying patience, and providing structure (e.g., rules, expectations, assessment criteria) in autonomy-supportive ways (Reeve & Cheon, 2021). Such a motivating style precluded students from (1) working against (conflict) and (2) striving to out-perform (competition/ability) their classmates, using the terminology and measures from Ntoumanis and Vazou (2005). When teachers move peer-to-peer interactions away from conflict, they in effect prevent the norms, expectancies, and patterns of communication that, if left unchecked, tend to breed antisocial behavior.

7.2. Needs are important too

While the "Needs and Climates" model better explained lesser antisocial behavior, needs are still important. Mid-semester (T2) increases in L1 need frustration did increase end-of-semester (T3) L1 antisocial behavior ($\beta = 0.17$, $p < .001$; see lower part of Fig. 5). This means that individual students who experienced a greater-than-class-average rise in their need frustration during PE instruction (for whatever reason) also reported a greater-than-class-average rise in their end-of-semester antisocial behavior. This same L1 effect has been reported in other investigations (Cheon et al., 2018; Pavey, Greitemeyer, & Sparks, 2011; Tian, Zhang, & Huebner, 2018), presumably because individually-experienced need frustration gives rise to frustration-infused negative emotions (e.g., anger), self-focused concerns, social dominance goals, and poor ways of coping (Delrue et al, 2017; McHoskey, 1999). The limitation of this effect, however, was that we delivered the intervention at the L2 (to teachers), not at the L1 (to the individual student), level. These effects therefore occurred outside of the intervention effect.

The teacher-delivered intervention did decrease class-wide need frustration ($\beta = -0.66$, $p < .001$; see upper part of Fig. 5). This is an important finding, because we suspect that decreased class-wide need frustration can explain many student outcomes, such as amotivation, negative emotionality, stress, burnout, and so forth. But the dependent measure in the present study was antisocial behavior. We selected this particular outcome because many educators consider antisocial behavior to be largely a group-generated and -regulated behavior (Hendrickx et al., 2016). Thus, we suspect that intervention-enabled changes in class-wide need satisfaction and frustration best explain some student outcomes (e.g., well-being), while intervention-enabled changes in the supportive and conflictual classroom climate best explain different student outcomes (e.g., antisocial behavior). We further suspect that a "Needs and Climates" model provides a more comprehensive explanatory model than does a "Needs Only" model. By "a more comprehensive explanatory model", we mean that a "Needs and Climates" model, compared to a "Needs Only" model, explains (1) a wider range of outcomes (i.e., expands the number of student outcomes explained) and (2) a greater proportion of the variance in any one particular outcome (i.e., the R^2 for L2 T3 antisocial behavior rose from 0.51 in the "Needs Only" model to 0.86 in the "Needs and Climates" model).

7.3. Limitations, a question, future research, and conclusion

Limitations. Two methodological decisions limited the conclusions that can be reached. First, we assessed antisocial behavior only at the classroom level (e.g., "In this class, my classmates …"). To build on the present findings, the research design could now be expanded by asking students to report *both* their individual experience (e.g., "In this class, I fight.") and their perception of the group reality (e.g., "In this class, my classmates fight."). Assessing dependent measures with questionnaires that use both a student referent and a classroom referent can (1) disentangle L1 individual difference variables (psychological needs) from L1 residual effects of aggregated L2 climates and contexts (classroom climates), (2) investigate both L1 individual processes and L2 climate processes as important outcomes, and (3) evaluate potential cross-level interactions between L1 individual differences and L2 climate effects (e.g., Marsh et al., 2011; Pagaioannou et al., 2004).

Second, the concepts and measures for need frustration, conflictual climate, and antisocial behavior overlapped considerably to produce some conceptual and assessment ambiguity. Some of this overlap can be traced to similar item content, so future studies could address this as a measurement issue. That said, we argue that these concepts can be distinguished in two crucial ways. First, psychological needs (relatedness frustration) are assessed with items using the student as the referent, while classroom climate and antisocial behavior are assessed with items using the classroom at the referent (i.e., they measure two different things). Second, the "how to" of reducing antisocial behavior can leave teachers bewildered (e.g., "How do I get my students to stop fighting?"). We suggest that autonomy-supportive teaching influences the *antecedent* of antisocial behavior (i.e., conflictual climate), but not necessarily antisocial behavior per se (other than through the indirect effect of an improved classroom climate). This is an important distinction because, while antisocial behavior can be difficult for teachers to influence, they can influence its malleable and controllable antecedent. We suggest teachers invest their efforts in diffusing and preventing a conflictual classroom climate (e.g., students blame and criticize each other).

A Question. The DL-ML-SEM model is a sophisticated analytic model ($N = 1487$ students nested within 49 classes) that raises the question of why we did not just employ a simpler class-average analytical model ($N = 49$ teachers/classrooms)? The reason is that a L2 class phenomenon has a different interpretation under a class-average model vs. a DL-ML-SEM model. In the class-average model, the L2 dependent measure is a class-average score—the average score of all the individual students in that class. In a DL-ML-SEM model, the L2 dependent measure is a collectively-experienced classroom climate score—the group consensus about how prevalent that dependent measure is perceived to be. In other words, the class-average model analyzes *M* scores, while the DL-ML-SEM analyzes the variance in student ratings captured by *ICC* scores (i.e., an emergent process from the L1 ratings and not just an average of those ratings).

Future Research. Past autonomy-supportive teaching interventions have relied on the student as the unit-of-analysis. This has worked well to advance the literature, but it underestimates the benefits of greater autonomy-supportive teaching. To better measure and understand L2 processes, we encourage future researchers to adopt the classroom as the unit-of-analysis (or both the student and the classroom as multiple units-of-analysis, using a DL-ML-SEM analytical framework) as they focus on additional student outcomes, such as class-wide engagement, skill development, and physical activity level.

Given these positive findings on reduced class-wide antisocial behavior, future research can now take the next step in this program of research to focus on enhanced class-wide prosocial behavior. Such a future study could employ the same theoretical model and utilize the same research design as in the present study. We suspect that teacher participation in an autonomy-supportive teaching intervention would both diminish antisocial behavior via its lessening effect on a conflictual classroom climate and enhance prosocial behavior via its favorable effect on a supportive classroom climate.

Teacher participation in the autonomy-supportive teaching workshop did increase class-wide need satisfaction (H1) and decrease class-wide need frustration (H2). While these changes in students' psychological needs did not explain independent variance in class-wide antisocial behavior, it is possible that these changes affect the emerging classroom climate. If so, future research could test a double mediation model in which experimental condition affects psychological needs, which affect classroom climates, which affect antisocial behavior (or experimental condition affects classroom climate, which affect psychological needs, which affect antisocial behavior).

Another promising next step in this area of research would be to use SDT-centric concepts and measures to represent the classroom climate, such as peer-to-peer autonomy support and peer-to-peer interpersonal control. Such measures do not yet exist, so we suggest that measures of autonomy-supportive and controlling teaching (that use the teacher as the item referent) could be successfully adapted to create parallel measures of autonomy-supportive and controlling classroom climates (that use one's classmates as the item referent).

Conclusion. We applied a classroom unit-of-analysis and a DL-ML-SEM framework to study the benefits of greater autonomy-supportive teaching. Doing so produced two conclusions: (1) the autonomy-supportive teaching intervention improved both student psychological needs and the classroom climate and (2) it was intervention-enabled changes in the classroom climate that best explained changes in antisocial behavior.

Conflicts of Interest

The authors declare no conflict of interest.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.psychsport.2022.102174.

References

deCharms, R. (1976). *Enhancing motivation: Change in the classroom*. New York: Irvington.

Abula, K., Beckmann, J., He, Z., Cheong, C., Lu, F., & Gropel, P. (2020). Autonomy support in physical education promotes autonomous motivation towards leisure-time physical activity: Evidence from a sample of Chinese college students. *Health Promotion International, 35*(1). https://doi.org/10.1093/heapro/day102. February, e1-e10.

Assor, A., Feinberg, O., Kanat-Maymon, Y., & Kaplan, H. (2018). Reducing violence in non-controlling ways: A change program based on self-determination theory. *The Journal of Experimental Education, 86*(2), 195–213. https://doi.org/10.1080/00220973.2016.1277336

Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., Bosch, J. A., & Thøgersen-Ntoumani, C. (2011a). Self-determination theory and diminished functioning: The role of interpersonal control and psychological need thwarting. *Personality and Social Psychology Bulletin, 37*, 1459–1473. https://doi.org/10.1177/0146167211413125

Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., & Thøgersen-Ntoumani, C. (2011b). Psychological need thwarting in the sport context: Assessing the darker side of athletic experience. *Journal of Sport & Exercise Psychology, 33*, 75–102.

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education, 84*, 740–756. https://doi.org/10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3

Boardley, I. D., & Kavussanu, M. (2010). Effects of goal orientation and perceived value of toughness on antisocial behavior in soccer: The mediating role of moral disengagement. *Journal of Sport & Exercise Psychology, 32*, 176–192.

Cheon, S. H., Kim, S. J., Song, Y. G., & Lee, C. (2017). Relations between teacher interpersonal communication styles and student prosocial-and-antisocial behaviors in physical education context. *Korean Journal of Physical Education, 54*(4), 175–190.

Cheon, S. H., & Reeve, J. (2013). Do the benefits from autonomy-supportive PE teacher training programs endure?: A one-year follow-up investigation. *Psychology of Sport and Exercise, 14*, 508–518. https://doi.org/10.1016/j.psychsport.2013.02.002

Cheon, S. H., & Reeve, J. (2015). A classroom-based intervention to help teachers decrease students' amotivation. *Contemporary Educational Psychology, 40*, 99–111.

Cheon, S. H., Reeve, J., Lee, Y., Ntoumanis, N., Gillet, N., Kim, B. R., & Song, Y.-G. (2019). Expanding autonomy psychological need states from two (satisfaction, frustration) to three (dissatisfaction): A classroom-based intervention study. *Journal of Educational Psychology, 111*, 685–702.

Cheon, S. H., Reeve, J., & Moon, I. S. (2012). Experimentally-based, longitudinally designed, teacher-focused intervention to help physical education teachers be more autonomy supportive toward their students. *Journal of Sport & Exercise Psychology, 34*, 365–396. https://doi.org/10.1123/jsep.34.3.365

Cheon, S. H., Reeve, J., & Ntoumanis, N. (2018). A needs-supportive intervention to help PE teachers enhance students' prosocial behavior and diminish antisocial behavior. *Psychology of Sport and Exercise, 35*, 74–88.

Cheon, S. H., Reeve, J., & Ntoumanis, N. (2019). An intervention to help teachers establish a prosocial peer climate in physical education. *Learning and Instruction, 64*, 1–14. https://doi.org/10.1016/j.learninstruc.2019.101223. Article 101223.

Cheon, S. H., Reeve, J., & Song, Y.-G. (2016). A teacher-focused intervention to decrease PE students' amotivation by increasing need satisfaction and decreasing need frustration. *Journal of Sport & Exercise Psychology, 38*, 217–235.

Cheon, S. H., Reeve, J., & Song, Y.-G. (2019). Recommending goals and supporting needs: An intervention to help physical education teachers communicate their expectations while supporting students' psychological needs. *Psychology of Sport and Exercise, 41*, 107–118. https://doi.org/10.1016/j.psychsport.2018.12.008

Cheon, S. H., Reeve, J., & Vansteenkiste, M. (2021). When teachers learn how to provide classroom structure in an autonomy-supportive way: Benefits to teachers and their students. Teaching and Teacher Education. Article 103004 https://doi.org/10.1016/j.tate.2019.103004.

Cheon, S. H., & Song, Y. G. (2011). Korean version of peer motivational climate in youth Sport Questionnaire: Development and validity. *Korean Journal of Sport Psychology, 22*(1), 243–260.

Cohen, J. (1988). In *Statistical power analysis for the behavioral sciences* (2e.). Hillsdale, NJ: Lawrence Erlbaum.

Cook, C. R., Williams, K. R., Guerra, N. G., Kim, T. E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly, 25*, 65–83.

Delrue, J., Vansteenkiste, M., Mouratidis, A., Gevaert, K., Vande Broek, G., & Haerens, L. (2017). A game-to-game investigation of the relation between need-supportive and need-thwarting coaching and moral behavior in social. *Psychology of Sport and Exercise, 31*, 1–10.

Espelage, D. L., Holt, M. K., & Henkel, R. R. (2003). Examination of peer-group contextual effects on aggression during early adolescence. *Child Development, 74*, 205–220.

Espelage, D. L., & Swearer, S. M. (Eds.). (2011). *Bullying in North American schools* (2nd ed.). New York, NY: Routledge.

Flunger, B., Mayer, A., & Umbach, N. (2019). Beneficial for some or for everyone? Exploring the effects of an autonomy-supportive intervention in the real-life classroom. *Journal of Educational Psychology, 111*(2), 210–234. https://doi.org/10.1037/edu0000284

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95*(1), 148–162.

Gunnell, K., Crocker, P. R. E., Wilson, P. M., Mack, D. E., & Zumbo, B. D. (2013). Psychological need satisfaction and thwarting: A test of basic psychological needs theory in physical activity contexts. *Psychology of Sport and Exercise, 14*, 599–607.

Haerens, L., Aelterman, N., Vansteenkiste, M., Soenens, B., & Van Petegem, S. (2015). Do perceived autonomy-supportive and controlling teaching relate to physical education students' motivational experiences through unique pathways? Distinguishing between the bright and dark side of motivation. *Psychology of Sport and Exercise, 16*, 26–36. https://doi.org/10.1016/j.psychsport.2014.08.013

Hendrickx, M. M. H. G., Mainhard, M. T., Boor-Klip, H. J., Cillessen, A. H. M., & Brekelmans, M. (2016). Social dynamics in the classroom: Teacher support and conflict and the peer ecology. *Teaching and Teacher Education, 53*, 30–40.

Hodge, K., & Gucciardi, D. F. (2015). Antisocial and prosocial behavior in sport: The role of motivational climate, basic psychological needs, and moral disengagement. *Journal of Sport & Exercise Psychology, 37*, 257–273.

Hong, J. S., & Espelage, D. L. (2012). A review of research on bullying and peer victimization in school: An ecological system analysis. *Aggression and Violent Behavior, 17*, 311–322. https://doi.org/10.1016/j.avb.2012.03.003

Jang, H., Kim, E.-J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction, 43*, 27–38.

Jang, H., Reeve, J., Ryan, R. M., & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically-oriented South Korean adolescents? *Journal of Educational Psychology, 101*, 644–661.

Joesaar, H., Hein, V., & Hagger, M. S. (2012). Youth athletes' perception of autonomy support from the coach, peer motivational climate and intrinsic motivation in sport setting: One-year effects. *Psychology of Sport and Exercise, 13*, 257–262.

Juvonen, J., & Graham, S. (2014). Bullying in schools: The power of bullies and the plight of victims. *Annual Review of Psychology, 65*, 159–185. https://doi.org/10.1146/annurevpsych-010213-115030

Kaplan, H., & Assor, A. (2012). Enhancing autonomy-supportive I-Thou dialogue in schools: Conceptualization and socio-emotional effects of an intervention program. *Social Psychology of Education, 15*, 251–269. https://doi.org/10.1007/s11218-012-9178-2

Karna, A., Voeten, M., Little, T., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large scale evaluation of the KiVa anti-bullying program: Grades 4-6. *Child Development, 82*, 311–320.

Kavussanu, M. (2012). Moral behavior in sport. In S. Murphy (Ed.), *The Oxford handbook of sport and performance psychology* (pp. 364–383). Oxford: Oxford University Press.

Klein, K. J., & Kozlowski, S. W. (2000). From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods, 3*, 211–236.

Langan, E., Toner, J., Blake, C., & Lonsdale, C. (2015). Testing the effects of a self-determined theory-based intervention with youth Gaelic football coaches on athlete motivation and burnout. *The Sport Psychologist, 29*(4), 293–301. https://doi.org/10.1123/TSP.2013-0107

Lansford, J. E., Malone, P. S., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2010). Developmental cascades of peer rejection, social information processing biases, and aggression during middle childhood. *Development and Psychopathology, 22*, 593–602.

Lonsdale, C., Rosenkranz, R. R., Sanders, T., Peralta, L. R., Bennie, A., Jackson, B., Taylor, I. M., & Lubans, D. R. (2013). A cluster randomized control trial of strategies to increase adolescents' physical activity and motivation in physical education: Results of the Motivating Active Learning in Physical Education (MALP) trial. *Preventive Medicine, 57*, 696–702. https://doi.org/10.1016/j.ypmed.2013.09.003

Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2x2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error-correction models. *Psychological Methods, 16*(4), 444–467. https://doi.org/10.1037/a0024376

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*, 203–229.

Manninen, M., Deng, Y., Hwang, Y., Waller, S., & Yli-Piipari, S. (2020). Psychological need-supportive instruction improves novel skill performance, intrinsic motivation, and enjoyment: A cluster-randomized trial. *International Journal of Sport and Exercise Psychology*. https://doi.org/10.1080/1612197X.2020.1826999

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares, & J. McCardie (Eds.), *Psychometrics: A. Festschrift to Roderick P. McDonald* (pp. 275–340). Hillsdale, NJ: Erlbaum.

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Koller, O. (2011). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*(2), 106–124. https://doi.org/10.1080/00461520.2012.670488

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthen, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*, 764–802.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise & Sport, 60*, 48–58.

McHoskey, J. W. (1999). Machiavellianism, intrinsic versus extrinsic goals, and social interest: A self-determination theory analysis. *Motivation and Emotion, 23*, 267–283.

Mikami, A. Y., Lerner, M. D., & Lun, J. (2010). Social context influences on children's rejection by their peers. *Child Development Perspectives, 4*, 123–130.

Morin, A. J. S., Blais, A.-R., & Chenard-Poirier, L. A. (2021). Doubly latent multivariate analyses procedures for organizational assessment and prediction. *Journal of Business and Psychology*. https://doi:10.1007/s10869-021-09736-5.

Morin, A. J. S., Marsh, H. W., Nagengast, B., & Scalas, L. F. (2014). Doubly latent multivariate analyses of classroom climate: An illustration. *The Journal of Experimental Education, 82*(2), 143–167. https://doi.org/10.1080/00220973.2013.769412

Muthén, L., & Muthén, B. (2019). *Mplus user's guide. Angeles: Muthén & Muthén*.

Ntoumanis, N., & Vazou, S. (2005). Peer motivational climate in youth sport: Measurement development and validation. *Journal of Sport & Exercise Psychology, 27*, 432–455.

Ntoumanis, N., Vazou, S., & Duda, J. (2007). Peer-created motivational climate. In S. Jowett, & D. Lavallee (Eds.), *Social psychology in sport* (pp. 145–156). Champaign, IL: Human Kinetics.

OECD. (2020). *Global teaching insights: A video study of teaching*. Paris: OECD Publishing. https://doi.org/10.1787/20d6f36b-en

Papaioannou, A., Marsh, H. W., & Theodorakis, Y. (2004). A multilevel approach to motivational climate in physical education and sport settings: An individual or a group level construct. *Journal of Sport & Exercise Psychology, 26*, 90–118.

Patall, E. A., Steingut, R. R., Vasquez, A. C., Trimble, S. R., Pitch, K. A., & Freeman, J. L. (2018). Daily autonomy supporting or thwarting and students' motivation, engagement, and disaffection in the high school science classroom. *Journal of Educational Psychology, 110*(2), 268–288. https://doi.org/10.1037/edu0000214

Pavey, L., Greitemeyer, T., & Sparks, P. (2011). Highlighting relatedness promotes prosocial motives and behavior. *Personality and Social Psychology Bulletin, 37*(7), 905–917.

Pellegrini, A. D. (2001). Affiliative and aggressive dimensions of dominance and possible functions during early adolescence. *Aggression and Violent Behavior, 7*, 21–31.

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures, 6*, 77–98.

Reeve, J., & Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist, 56*(1), 54–77. https://doi.org/10.1080/00461520.2020.1862657

Reeve, J., Cheon, S. H., & Yu, T. H. (2020). An autonomy-supportive intervention to develop students' resilience by boosting agentic engagement. *International Journal of Behavioral Development, 44*(4), 325–338. https://doi.org/10.1177/0165025420911103

Reeve, J., Ryan, R. M., Cheon, S. H., Matos, L., & Kaplan, H. (2022). *Supporting students' motivation: Strategies for success*. London, UK: Routledge.

Richey, J. E., Bernacki, M. L., Belenky, D. M., & Nokes-Malach, T. J. (2018). Comparing class- and task-level measures of achievement goals. *The Journal of Experimental Education, 86*(4), 560–578. https://doi.org/10.1080/00220973.2017.1386155

Roth, G., Kanat-Maymon, Y., & Bibi, U. (2010). Prevention of school bullying: The important role of autonomy-supportive teaching and internalization of prosocial values. *British Journal of Educational Psychology, 81*(4), 654–666. https://doi.org/10.1348/2044-8279.0020003

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Press.

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15*(2), 112–120. https://doi.org/10.1016/j.avb.2009.08.007

Selig, J. P., & Preacher, K. J. (2008). Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects [Computer software]. Available from: http://quantpsy.org/.

Shields, D. L., Funk, C. D., & Bredemeier, B. L. (2015). Predictors of moral disengagement in sport. *Journal of Sport & Exercise Psychology, 37*, 646–658.

Smith, A. L. (2003). Peer relationships in physical activity contexts: A road less traveled in youth sport and exercise psychology research. *Psychology of Sport and Exercise, 4*, 25–39.

Standage, M., Duda, J. L., & Ntoumanis, N. (2006). Students' motivational processes and their relationship to teacher ratings in school physical education: A self-determination theory approach. *Research Quarterly for Exercise & Sport, 77*, 100–110.

Thomas, D. E., Bierman, K. L., et al. (2006). The impact of classroom aggression on the development of aggressive behavior problems in children. *Development and Psychopathology, 18*, 471–487.

Thornberg, R., Wanstrom, L., & Jungert, T. (2018). Authoritative classroom climate and its relations to bullying victimization and bystander behaviors. *School Psychology International, 39*(6), 663–680. https://doi.org/10.1177/0143034318809762

Tian, L., Zhang, X., & Huebner, E. S. (2018). The effect of satisfaction of basic psychological needs at school on children's prosocial behavior and antisocial behavior: The mediating role of school satisfaction. *Frontiers in Psychology: Educational Psychology, 9*, 548. https://doi.org/10.3389/fpsyg.2018.00548

Tilga, H., Hein, V., & Koka, A. (2019). Effects of a web-based intervention for PE teachers on students' perceptions of teacher behaviors, psychological needs, and intrinsic motivation. *Perceptual & Motor Skills, 126*(3), 559–580. https://doi.org/10.1177/0031512519840150

Ulstad, S. O., Halvari, H., Sorebo, O., & Deci, E. L. (2018). Motivational predictors of learning strategies, participation, exertion, and performance in physical education: A randomized control trial. *Motivation and Emotion, 42*(4), 497–512. https://doi.org/10.1007/s11031-018-9694-2

Van Petegem, S., Soenens, B., Vansteenkiste, M., & Beyers, W. (2015). Rebels with a cause? Adolescent defiance from the perspective of reactance theory and self-determination theory. *Child Development, 86*(3), 903–918. https://doi.org/10.1111/cdev.12355

Vazou, S., Ntoumanis, N., & Duda, J. L. (2006). Predicting young athletes' motivational indices as a function of their perceptions of the coach- and peer-created climate. *Psychology of Sport and Exercise, 7*, 215–233.