Journal of School Psychology 110 (2025) 101431

Contents lists available at ScienceDirect

Journal of School Psychology

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

Bystanders tend to defend victims in a supportive classroom climate: A cluster randomized control trial and an observational study

Sung Hyeon Cheon [a,*], Johnmarshall Reeve [b,**], Kyoung-Eun Yoo [a], Yong-Gwan Song [c], Herbert W. Marsh [b], Hye-Ryen Jang [b], Youngsun Lee [d]

[a] Department of Physical Education, Korea University, South Korea
[b] Institute for Positive Psychology and Education, Australian Catholic University, Australia
[c] Division of Smart Healthcare, Pukyong National University, South Korea
[d] Department of Special Education, Ewha Womans University, South Korea

ARTICLE INFO

Editor: Andy Garbacz

Keywords:
Autonomy Support
Bystander
Classroom climate
Defending
Social Ecology

ABSTRACT

When bystanders reinforce bullies, bullying tends to escalate; when bystanders defend victims, bullying tends to de-escalate. Recognizing this pattern, we adopted a social-ecological perspective within a self-determination theory framework to conduct two studies. Study 1 was a pre-registered cluster randomized control trial in which 38 Korean secondary teachers (9 females, 29 males; 19 experimental, 19 control) participated in an intervention to learn how to teach in highly autonomy-supportive and not-at-all controlling ways. We hypothesized that this approach to teaching would create a supportive peer-to-peer classroom climate, which in turn would increase defending- and decrease passive- and reinforcing-bystanding. In three waves over an academic year, 1084 adolescent students (490 females, 594 males) reported their classroom climate and bystanding behaviors. According to a doubly latent multilevel SEM analysis, experimental group teachers created a more supportive climate ($B = 0.55$, $p < .001$) and this classroom-level effect increased defending-bystanding ($B = 0.55$, $p = .001$), decreased passive-bystanding ($B = -0.52$, $p < .001$), and decreased reinforcing-bystanding ($B = -0.40$, $p = .006$). Study 2 was a correlational study in which 629 adolescent students (398 females, 231 males) reported on their teacher's autonomy-supportive and controlling motivating styles and the 11 teachers (four females, seven males) rated each student on the three bystanding behaviors and the extent to which they contributed to two dimensions of classroom climate (i.e., supportive and conflictual). A SEM analysis showed that students' perceived autonomy-supportive teaching predicted teacher-rated supportive climate ($B = 0.23$, $p = .036$) and students' perceived controlling teaching predicted teacher-rated conflictual climate ($B = 0.11$, $p = .041$). According to mediation analyses, supportive climate ratings then predicted teacher-rated high defending ($B = 0.28$, $p = .006$) and low passive ($B = -0.29$, $p < .001$) bystander behavior, whereas conflictual climate ratings predicted teacher-rated low defending ($B = -0.22$, $p = .008$) and high passive ($B = 0.26$, $p = .001$) and high reinforcing ($B = 0.37$, $p < .001$) bystander behavior. We conclude that teachers can learn how to

* Correspondence to: S. H. Cheon, Department of Physical Education, Korea University, Anam-dong, Seongbuk-gu, Seoul 02841, South Korea.
** Correspondence to: J. Reeve, Institute for Positive Psychology and Education, Australian Catholic University, North Sydney Campus, 33 Berry Street, 9th floor, Sydney 2060, Australia.
E-mail addresses: cheon78@korea.ac.kr (S.H. Cheon), johnmarshall.reeve@acu.edu.au (J. Reeve).

https://doi.org/10.1016/j.jsp.2025.101431
Received 8 September 2023; Received in revised form 17 January 2025; Accepted 23 January 2025
Available online 28 March 2025

create a supportive climate that orients students toward defending and away from passive and reinforcing bystanding.

1. Introduction

Bystanders play a crucial role in escalating versus de-escalating peer victimization (Hong & Espelage, 2012; Salmivalli, 2014). For instance, in the presence of reinforcing bystanders, bullying tends to escalate, but in the presence of defending bystanders, it tends to de-escalate (Kärnä et al., 2010, 2011; Salmivalli et al., 2011). Such findings put an explanatory spotlight on the potentially important role bystanders may play in the group dynamics of peer bullying (Polanin et al., 2012; Salmivalli, 2014).

Three bystander behaviors are common (Jungert et al., 2016; Poyhonen et al., 2010; Pozzoli et al., 2012; Salmivalli, 2014; Salmivalli et al., 2011). The first is *reinforcing*, also referred to as "assistant" (Salmivalli, 2014) or "bully empowerment" (Marsh et al., 2023). In this role, the bystander amplifies bullying by coming over to side with the bully, providing an encouraging audience, reinforcing the bullying by laughing, smiling, cheering, or gesturing enthusiastically, or joining in on the bullying. The second is the *passive bystander*, also referred to as "outsider" (Thornberg et al., 2012) or "ignore the situation" (Marsh et al., 2023). In this role, the bystander remains uninvolved by not doing anything, simply staying away, or walking away from and ignoring the bullying episode. The third is *defending* (Lambe et al., 2019), also referred to as "victim empowerment" (Marsh et al., 2023). In this role, the bystander de-escalates bullying by coming over to side with the victim, saying "no" to the bully, informing a teacher, helping the victim get away, or by listening to, talking with, advising, or comforting the victim.

2. Social-ecological model to explain bystander behavior

A social-ecological model highlights the important role that the peer culture plays in understanding bullying (Hong & Espelage, 2012; Salmivalli, 2010). This model views bullying as a social phenomenon regulated by group norms, status concerns, and social contextual forces, rather than as an individual phenomenon in which a few "bad" kids attack a few "weak" kids (Espelage & Swearer, 2004). According to this model, peer bystander behaviors are central to establishing, maintaining, preventing, escalating, and de-escalating peer victimization (Hong & Espelage, 2012; Salmivalli et al., 2011; Van Ryzin & Roseth, 2018).

When classroom-based (in Korea, Grades 7–12; Cheon, Reeve, Marsh, & Jang, 2023) and school-based (in Finland, Grades 4–6; Kärnä et al., 2011) interventions have systematically changed peer bystander behavior to increase defending, peer victimization has declined (see Polanin et al., 2012, for a meta-analysis of bystander interventions). Within the context of an intervention, students can learn effective ways of defending and, when enacted competently and with the support of one's peers, defending tends to be associated with reduced bullying. However, outside of such an intervention, defending tends not to reduce bullying (Gaffney et al., 2021). In addition, in a typical classroom or school situation, bystanders rarely take the initiative to intervene to support the victim (Hawkins et al., 2001). Research explains why this might be, as defending can provoke retaliation from the bully and reinforcing bystanders and thus increase the defender's potential vulnerability to negative personal consequences, such as social isolation and anxiety (Hinduja & Patchin, 2010; Wu et al., 2016). Other research suggests that defending can lead to positive personal consequences (e.g., peer popularity), although this effect was observed only for girls (not for boys) and when those girls could expect peer support rather than peer retaliation (Laninga-Wijnen et al., 2023).

A recent investigation involving Finnish adolescents helps resolve this question of whether defending is effective and associated with positive repercussions or is ineffective and associated with negative repercussions (Malamut et al., 2023). These researchers showed that a key moderating variable may be the quality of the classroom climate: When prevailing classroom norms allocated high popularity to bullies, defending tended to be ineffective; but when prevailing classroom norms allocated high popularity to defenders, defending tended to be effective. What this means is that defending may need a precondition. Before encouraging adolescents to put themselves at risk by defending, we propose that the prevailing classroom climate likely needs to make it clear and consensual that one's peers will back and support their defending (Cheon, Reeve, Marsh, & Jang, 2023; Flaspohler et al., 2009). We further suggest that the classroom teacher can create such a supportive peer climate, but only if that teacher is first provided with professional training in how to do so.

3. Supportive classroom climate

Any classroom exudes a social climate. That climate can range from one that is hierarchical and competitive to one that is egalitarian and supportive (Garandeau et al., 2014). Status-centric, hierarchical climates tend to cultivate "me vs. you" peer interactions and social relationships. These social comparison-based interpersonal dynamics leave students vulnerable to status concerns and dominance hierarchies (i.e., "competition for social dominance"; Nocentini et al., 2013, p. 495) that can increase the likelihood of conflict (Di Stasio et al., 2016) and feed into the establishment and escalation of a competitive and conflictual classroom climate (Ntoumanis & Vazou, 2005).

Conversely, closely-knit and emotionally caring climates tend to cultivate a sense of community that encourages cooperation. These egalitarian-based interpersonal dynamics encourage students to connect with and support each other (Assor et al., 2018; Kaplan & Assor, 2012; Van Ryzin & Roseth, 2018) and can feed into a cooperative, supportive, and caring classroom climate (Ntoumanis & Vazou, 2005; Pozzoli et al., 2012). As shown in prior research, teachers can learn how to create such a closely-knit classroom climate,

such as through cooperative learning (Van Ryzin & Roseth, 2018) or autonomy-supportive teaching (Cheon, Reeve, Marsh, & Jang, 2023).

4. Autonomy-supportive teaching

Autonomy support is a construct derived from self-determination theory (Ryan & Deci, 2017). When applied to teaching, autonomy-supportive teaching is an approach to instruction and teacher-student relationships that allows students to satisfy their basic psychological needs (e.g., for autonomy) as they engage in classroom activities (Aelterman et al., 2019; Patall et al., 2013). When autonomy supportive, teachers rely on teaching practices such as taking the students' perspective, supporting students' interests and goals, and supporting students' volitional internalization of external regulations (Reeve et al., 2022; Reeve & Cheon, 2021). Autonomy-supportive teaching stands in contrast to controlling teaching. When controlling, teachers tell students what to do, what to think, and how to feel and they apply pressure or conditional regard until students comply with their prescriptions and proscriptions, irrespective of what students prefer (Soenens et al., 2012). When controlling, teachers rely on teaching practices such as uttering directives, using pressuring language, suppressing student input, and the giving and the taking away of environmental contingencies, such as incentives, punishers, and types of conditional regard (De Meyer et al., 2016).

Since Deci et al.' (1981) pioneering study, autonomy-supportive teaching has long been understood as an environmental-facilitating condition to enhance an individual student's motivation, engagement, learning, and prosocial behavior (Reeve et al., 2022). Recently, autonomy-supportive teaching interventions have expanded their scope beyond supporting the individual student's sense of autonomy to also supporting the whole class's sense of autonomy (Assor et al., 2018; Kaplan & Assor, 2012; Reeve et al., 2022). Autonomy-supportive teachers do this by relying on an empathic, considerate, and responsive communication and interaction style that helps students accept and volitionally internalize recommendations such as "be considerate", "value caring", "build friendships", and "show mutual respect" (Joussemet & Grolnick, 2022; Vansteenkiste et al., 2018). Because of this, we hypothesized that (a) teachers trained to be autonomy supportive and not-at-all controlling would be able to create a more supportive peer-to-peer classroom climate, and (b) this more supportive classroom climate, once established, would increase defending and decrease reinforcing- and passive-bystanding.

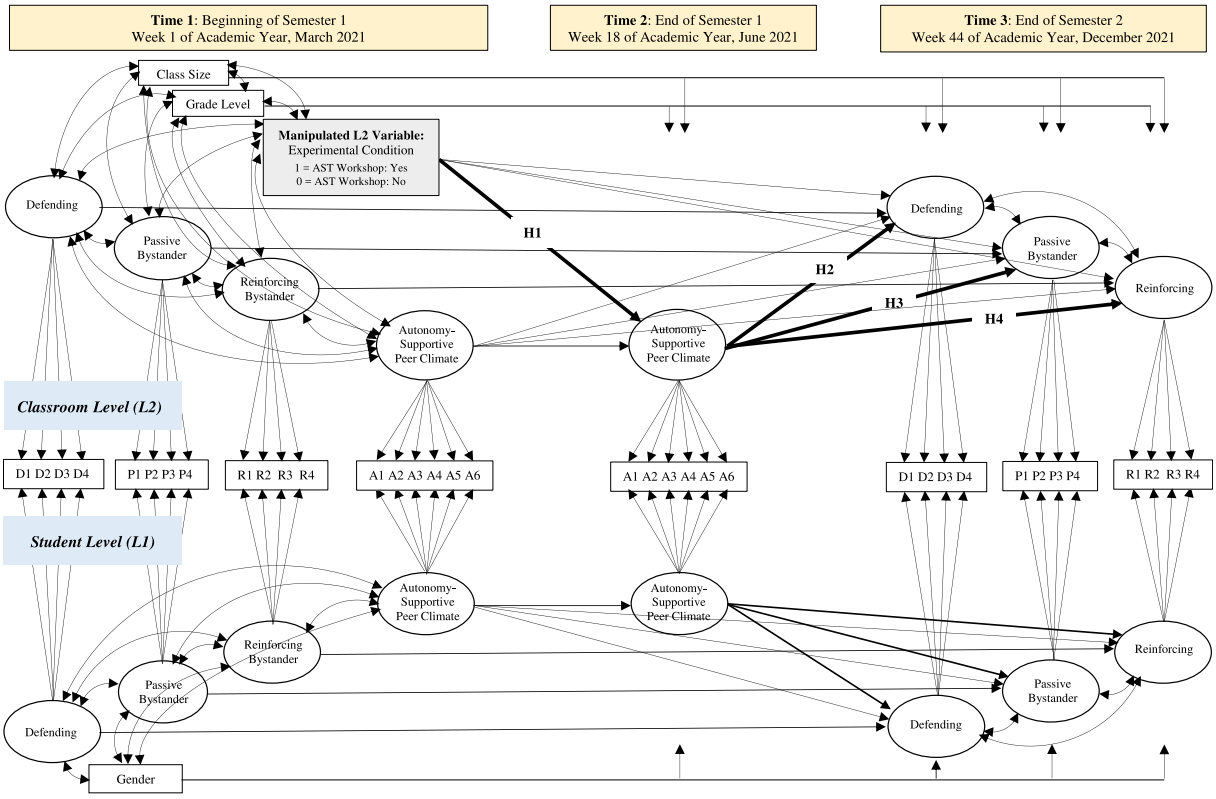

Fig. 1. DL-ML-SEM Hypothesized Model (Study 1).
Note. H = Hypothesis; D = Defending Bystander item; P = Passive Bystander item; R = Reinforcing Bystander item; A = Perceived Autonomy-Supportive Classroom Climate item. Thick lines are hypothesized paths. Thin lines are statistical controls, autoregressive effects, and indicators of a latent construct. Rectangles denote observed variables; ovals represent latent variables.

5. The present investigation

Study 1 employed a methodologically rigorous randomized control trial research design. Its purpose was to experimentally manipulate autonomy-supportive teaching to a high level and controlling teaching to a low level to test the intervention's capacity to create a more supportive peer-to-peer classroom climate that encourages class-wide levels of greater defending and lesser passive and reinforcing bystanding. Study 2 was a correlational study in which students reported their mid-semester perceptions of their teacher as autonomy supportive and as controlling while teachers rated their students' end-semester in-class displays of peer support, peer conflict, and the three bystanding behaviors. Study 2 added to Study 1 in two ways. First, Study 2 used teacher ratings (rather than student self-reports) to assess classroom climate and bystanding behavior. Second, Study 2 tested for separate effects from both autonomy-supportive and controlling teaching (not just autonomy-supportive teaching, as in Study 1) and it tested for separate effects from both a supportive and conflictual climate on the three bystanding behaviors (not just a supportive climate, as in Study 1). Together, the two studies allowed us to evaluate an overall hypothesized model in which high autonomy-supportive teaching and low controlling teaching would create a classroom climate that could encourage defending and discourage passive and reinforcing bystanding.

Both studies were conducted in Korean secondary grade level physical education (PE) classrooms. We focused on secondary grade classrooms because bullying tends to peak during the middle school years (ages 12–15 years; Hymel & Swearer, 2015). We focused on PE classrooms because PE settings (i.e., gym, sports field, and classroom) afford recurring opportunities for peer interaction and group dynamics. During PE activities (e.g., basketball, rope jumping), students have opportunities to interact with their peers in ways that are supportive (i.e., cooperative) or conflictual (i.e., competitive), including instances of bystanding during bully-victim episodes (Jiménez-Barbero et al., 2020). We focused on Korea because the Korean Ministry of Education has recognized bullying as a national concern with a high growth rate (Kim et al., 2019; Ministry of Education, 2021). Bullying's relatively frequent occurrence in Korea may be traced to somewhat unique societal forces, such as in-group versus out-group collectivistic social dynamics and Confucian beliefs that normalize interaction patterns such that younger students must respect older students.

6. Study 1

The hypothesized model and its four embedded hypotheses appear in Fig. 1. Specifically, we hypothesized that teacher-participants in the experimental condition, compared to teacher-participants in the control group, would create a supportive classroom climate during Semester 1 (Hypothesis 1). This more supportive classroom climate would then longitudinally increase students' year-end defending (Hypothesis 2) and longitudinally decrease year-end passive- (Hypothesis 3) and reinforcing- (Hypothesis 4) bystanding. Hypotheses 2, 3, and 4 represent a mediation model (i.e., experimental condition ➔ classroom climate ➔ bystander behavior), so we included follow-up tests for mediation.

7. Method

7.1. Transparency and openness

We preregistered this study with the Open Science Framework (OSF); in addition to the pre-registration document, the OSF materials include the study questionnaire, dataset, Mplus syntax used in all analyses, the CONSORT Checklist (for a cluster randomized control trial), a step-by-step description of the autonomy-supportive teaching workshop, and the observational rating sheet teachers used in Study 2.[1]

7.2. Participants

Thirty-eight PE teachers (9 women, 29 men) from 38 different schools (26 middle, 12 high schools) in Seoul, Korea served as participants. Teachers were, on average, 36.3 years old (SD = 5.2, range = 25–44) with 8.6 years (SD = 4.0, range = 1–14) of PE teaching experience. We recruited this sample of teachers from a larger pool of secondary grade PE teachers who had earlier attended a workshop on a different topic hosted by the Korean Ministry of Education. We recruited only one teacher from a school so that we could randomly assign teachers to the experimental condition without a risk of intra-school cross-condition contamination. Each teacher taught in a middle-class, neighborhood-based, medium-sized public school in the Seoul metropolitan area. PE classes were taught three times per week and each class lasted 50–55 min.

The students of these teachers included 1084 middle-class Korean adolescents (M_{age} = 14.3 years, SD = 1.5, range = 11–18), including 490 (45.2 %) females and 594 (54.8 %) males, 714 (65.9 %) middle and 370 (34.1 %) high schoolers, and 554 (51.1 %) students in the experimental group and 530 (48.9 %) in the control group. As to a priori sample size adequacy, this sample size (38 L2 units/teachers, 1084 L1 units/students) generally met the guidelines recommended for a multilevel structural equation modeling analysis (Lüdtke et al., 2011; Morin et al., 2021).

[1] See the OSF project site: https://osf.io/4akns/?view_only=88d78e48fc06412f928642c6d84f0d27

7.3. Multilevel research design

By adopting a social-ecological perspective, we felt it was necessary to test our hypothesized model with a multilevel perspective—one that can distinguish between individual student (L1) vs. classroom level (L2) effects (Cheon, Reeve, Marsh, & Jang, 2023; Pozzoli et al., 2012) —because bullying largely revolves around social or group-based (L2) processes. To the extent that this may be true, then any intervention designed to mobilize a supportive classroom climate and defending bystanders must start at the classroom level. Accordingly, we made a pre-registration decision to test our model (and all its hypotheses) at the classroom (L2) level.

We evaluated our hypothesized model by employing an experimental (intervention) research design with three waves of longitudinally assessed dependent measures over the course of an academic year. This experimental design (i.e., a randomized control trial) allowed us to assess the causal impact of manipulated autonomy-supportive teaching on a supportive classroom climate. The longitudinal data collection allowed us to investigate for changes in students' bystander behaviors over the course of the year. Of note, our preferred outcome measures were bystanding behaviors, rather than bullying or victimization per se. We made this pre-registration decision because, at the level of the individual student (L1), it is possible to distinguish between one student who is bullying and another who is bystanding: One student insults and hits, whereas the other observes and laughs. However, at the classroom level (L2), this distinction breaks down because reinforcing-bystanding is the same as bullying. For instance, one item on the Reinforcing Bystanding questionnaire is "I join in with the kids who call another kid nasty nicknames, threaten, hit, or push him/her" (Jungert et al., 2016, p. 83). In support of this reasoning, in a pair of previous investigations we found that bullying and reinforcing-bystanding correlated in the $r = 0.60$ range at the L1 individual student level but in the $r = 0.90$ range at the L2 group-based classroom level (Cheon et al., 2022; Cheon, Reeve, & Marsh, 2023). Presumably, this high L2 correlation occurs because bullying and reinforcing combine to produce a "mobbing" effect at the L2 or classroom level (Olweus, 1973). Therefore, to avoid introducing overwhelming

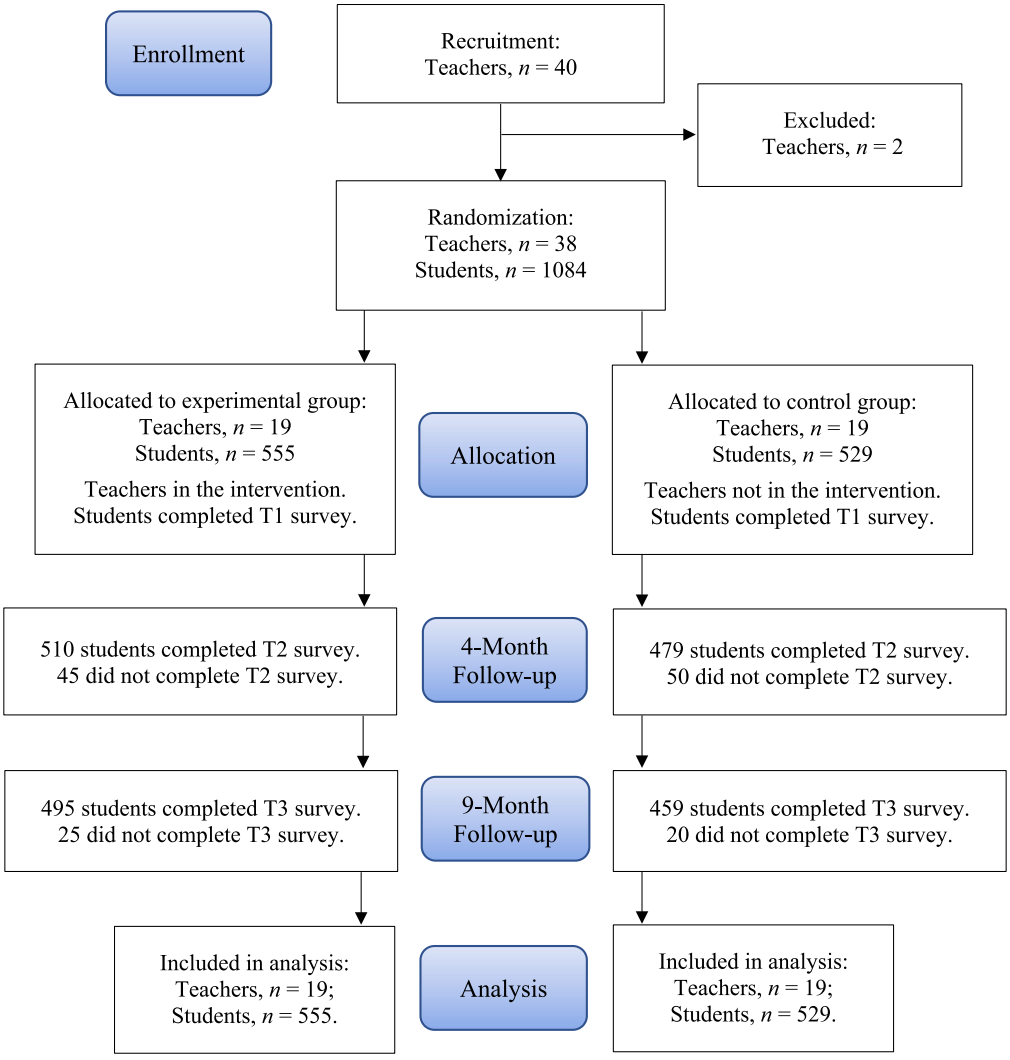

Fig. 2. Intervention flowchart for Study 1 (from CONSORT guidelines).

multicollinearity into the test of our hypothesized model (between L2 reinforcing and L2 bullying), we pre-registered the three bystander behaviors as our set of outcome measures.

7.4. Procedure

The research protocol was approved by Korea University's ethics committee. The research design was a cluster randomized control trial with longitudinally assessed dependent measures. Fig. 2 provides an overview and timeline for teacher recruitment (February 2021) and student data collection (March to December 2021). We randomly assigned teachers to condition using an online computer program. Teachers in the experimental condition participated in the Autonomy-Supportive Teaching (AST) workshop during the week before classes began for the academic year (February 2021). Data collection began in the first week of classes (first week in March 2021). Because it was the first week of classes, we did not expect the experimental condition to affect students' T1 (baseline) scores. Instead, these baseline scores served to represent students' early-course expectancies based on their past experiences with PE teachers and the PE course, their current PE teacher's reputation, past interactions with their classmates, and personal history with the sports, health, and exercise activities featured in the class. In our investigation, we were interested in how the teacher's experimental condition would affect a change in students' T2 and T3 scores from their T1 baseline levels. The T2 data collection occurred at the end of the first semester (June 2021), whereas the T3 data collection occurred at the end of the second semester (December 2021). For these data collection periods, we followed the schools' preference, request, and policy to use passive parental consent.

The 2021 academic year occurred during the Covid-19 pandemic. During the first semester, the Korean Ministry of Education established a grade-level rotation schedule in which the even grades attended face-to-face classes while the odd grades attended online classes. Then, during the next week, the odd grades attended face-to-face classes while the even grades attended online classes. During face-to-face classes, all students wore a facemask and social distancing was implemented. During the second semester, the grade level rotation schedule became optional so that each school made its own decision as to whether to return to all face-to-face classes, which many schools did.

Teachers in the experimental condition participated in a three-part, 8-h autonomy-supportive workshop (described below and in the Supplementary Materials), whereas teachers in the control condition were placed on a waitlist and completed the workshop at the beginning of the following year. For student data collection, the questionnaire began with a consent form. Students completed the questionnaire specifically in reference to their experience in the PE class. We assured students that their survey responses would be used only for the research study. Students in the experimental condition were not told that their teacher participated in the AST workshop.

7.5. Autonomy-Supportive Teaching (AST) workshop

The autonomy-supportive teaching (AST) workshop is based on the principles and applications of self-determination theory (SDT; Ryan & Deci, 2017). In the present study, its delivery followed the contents, activities, and step-by-step procedures of previously published AST workshops (e.g., Cheon et al., 2018; Cheon, Reeve, Marsh, & Jang, 2023). The Supplementary Materials provide step-by-step description of the AST workshop. In brief, Part 1 of the AST workshop was a 3-h, information-based morning presentation in which teachers learned about autonomy-supportive teaching, its benefits, its empirical evidence, its recommended practices, and its contrast with controlling teaching. Part 2 was a 3-h, skill-based afternoon workshop that took place on the same day as Part 1. It focused on the practical "how to" of the recommended autonomy-supportive instructional behaviors (ASIBs) by offering demonstrational examples, video modeling, step-by-step guidance, teaching simulations, and opportunities for discussion and reflection. During Part 2, teachers practiced, refined, and personalized the recommended teaching practices until they felt sufficiently willing and able to try them out in their own classrooms. For instance, two of the recommended ASIBs were (a) "take the students' perspective" (e.g., conduct a formative assessment [e.g., Q&A, use Mentimeter.com]) to become aware of what students are currently thinking and feeling about an assignment and (b) "provide an explanatory rationale" for each teacher request (e.g., "Let's use respectful language, *because* we want a classroom where everyone feels welcomed and safe"). Teachers also learned how to transform their existing controlling instructional behaviors into autonomy-supportive alternatives or substitutes (e.g., replace "criticize students' complaining" with "acknowledge and accept negative feelings"). Part 3 occurred a month into the first semester, which was after teachers had actual classroom experience with autonomy-supportive teaching. Part 3 featured a peer-to-peer group discussion to share teachers' early-semester experiences with autonomy-supportive teaching and to exchange ideas about how they might improve its classroom application.

Throughout the AST workshop, the focus was on becoming more autonomy supportive and less controlling during classroom instruction. No mention was made to bullying, victimization, or bystanding unless a participating teacher asked specifically about such a classroom situation. In past studies, teachers have reported high social fidelity following their AST workshop experience (> 6.0 on a 1–7 unipolar scale where $1 =$ *not at all* and $7 =$ *extremely*), reporting that they found the workshop experience to be useful, easily understandable, a catalyst to teaching effectiveness, and an experience they would recommend to their fellow teachers (Cheon et al., 2018; Cheon et al., 2024). To check for the successful implementation of the AST workshop in the present study, we collected a set of three manipulation checks. A manipulation check is a measure used to find out whether an experimental manipulation successfully delivered its intended effect, which in the present study was for teachers to provide instruction in a way that was high in autonomy support and low in teacher control. In the first manipulation check, students reported their perceived autonomy-supportive teaching on a questionnaire. In the second and third manipulation checks, classroom observers from the research team rated each teacher's in-class usage of autonomy-supportive and controlling teaching practices.

7.6. Raters' scoring of teachers' instructional behavior

Trained undergraduate and graduate research team members scored teachers' in-class instructional behaviors using the Behavior Rating Scale (BRS; Cheon et al., 2018) with a 7-point unipolar response scale (1 = *not at all*, 7 = *very much*). The BRS features six autonomy-supportive instructional behaviors (ASIBs; e.g., "takes the students' perspective") and six controlling instructional behaviors (CIBs; e.g., "uses pressuring language"). During Week 10 or 11, a pair of raters (blind to condition) visited all 38 teachers' classrooms, arrived 5–10 min before the start of class, and made independent ratings. All teachers knew in advance that the raters would visit their class once during the semester; however, we had raters show up unannounced to avoid any tendency for teachers to prepare that class differently from their other classes. The classroom observations did not occur until Week 10 or 11 to give teachers the time and experience they needed try out, refine, and personalize the recommended ASIBs. To estimate interrater reliabilities, we calculated the following statistics: (a) ICC for interrater consistency, (b) *r* for interrater correlation, and (c) kappa for interrater exact agreement. Raters' scorings were reasonably consistent across the six ASIBs (M_{ICC} = 0.85 [range = 0.80–0.89]; M_r = 0.74 [range = 0.66–0.82]; M_{kappa} = 0.44 [range = 0.17–0.63]) and across the six CIBs (M_{ICC} = 0.76 [range = 0.66–0.84]; M_r = 0.62 [range = 0.49–0.72]; M_{kappa} = 0.40 [range = 0.26–0.47]). To create the two manipulation check measures, we first averaged the two raters scores on each of these 12 instructional behaviors and then averaged the six ASIBs into one manipulation check score (six ASIBs ratings, ω = 0.95) and the six CIBs into a second manipulation check score (six CIB ratings, ω = 0.94).

7.7. Student questionnaire measures

All questionnaires were translated into Korean. We provide the full English version on the OSF project site. Each measure was originally developed in English, but we had available a previously utilized and professionally translated and back-translated Korean version (Cheon, Reeve, Marsh, & Jang, 2023). Questionnaires used a 7-point bipolar response scale (1 = *strongly disagree*, 7 = *strongly agree*). For each measure, we calculated the inter-item (i.e., McDonalds' omega [ω]) and inter-rater (ICC1) reliability statistics. The ICC1 communicates the proportion of variance in the dependent measure attributable to class membership. An ICC1 value >0.100 suggests a reasonable student-to-student consensual agreement on the prevailing level of that dependent measure in that particular class (Lüdtke et al., 2008; Morin et al., 2021).

Perceived Autonomy Support. For perceived autonomy support, we used two versions (i.e., teacher and peer) of the Learning Climate Questionnaire (LCQ; Black & Deci, 2000). "My teacher" was the item referent for the teacher version of the LCQ (e.g., "My PE teacher listens to how I would like to do things; six items; T1, T2, and T3 ωs were .89, .93, and .94, respectively). The purpose of this perceived autonomy-supportive teaching measure was to serve as a manipulation check as to whether teachers who participated in the AST workshop taught in a more autonomy supportive way than did teachers in the control group. "My classmates" was the item referent for the peer version of the LCQ (e.g., "My PE classmates try to understand how I see things before suggesting a new way to do things"; six items; T1, T2, and T3 ωs were 0.93, 0.95, and 0.95, respectively). The purpose of this perceived autonomy-supportive peers measure was to serve as the key mediating process in our hypothesized model (see the latent variable 'Autonomy-Supportive Peer Climate' in Fig. 1). By using these "My teacher" and "My classmates" referents, we assessed these two measures at the L2 level to represent the class consensus or "group-level reality" as to the prevailing level of that classroom phenomenon (Morin et al., 2022, p. 11). In confirmation that such an L2 group-reality did emerge in our investigation, students' L2 intraclass agreement on both measures was high (i.e., > 0.100) across all three waves of data collection, including autonomy-supportive teaching (ICC1s = 0.185, 0.174, and 0.200, respectively) and autonomy-supportive peer climate (ICC1s = 0.120, 0.141, and 0.164, respectively).

Bystander Measures. We used the Bystander Behavior Scale (BBS; Jungert et al., 2016) to measure defending, passive-bystanding, and reinforcing-bystanding.

Defending. We assessed defending with the four-item Defending scale from the BBS (e.g., "I do something to help if I see a kid being called nasty names, threatened, hit or pushed by other students"; Jungert et al., 2016; α = 0.70). In the present study, scores on the T1, T2, and T3 Defending scale were internally consistent (ωs = 0.89, 0.93, and 0.94, respectively), although only the T3 assessment showed a high within-class consensus (ICC1s = 0.064, 0.086, and 0.105, respectively).

Passive. We assessed passive-bystanding with the Passive Bystanding scale from the BBS (e.g., "I just ignore or no pay attention when I see my classmates bully someone"; Jungert et al., 2016). Because this three-item scale showed relatively low reliability in previous research (α = 0.57; Jungert et al., 2016), we made a pre-registration decision to create and add a fourth item: "When I see a student getting bullied, I just ignore or pay no attention to it." In the present study, scores on the resulting four-item Passive Bystanding scale were internally consistent for T1–T3 (ωs = 0.93, 0.95, and 0.93, respectively) and showed high within-class consensus (ICC1s = 0.168, 0.157, and 0.149, respectively).

Reinforcing. We assessed reinforcing-bystanding with the four-item Pro-Bullying scale from the BBS (e.g., "I laugh or cheer on the students who tease or call a classmate nasty nicknames"; Jungert et al., 2016; α = 0.73). In the present study, scores on the Pro-Bullying scale were internally consistent for T1–T3 (ωs = 0.93, 0.95, and 0.95, respectively) and showed a high within-class consensus (ICC1s = 0.273, 0.231, and 0.152, respectively).

7.8. Data analyses

Test of the Hypothesized Model. The data had a two-level structure with students (Level 1, N = 1084) nested within classrooms (Level 2, k = 38). We tested the overall hypothesized model depicted in Fig. 1 with a doubly latent multilevel structural equation model analysis (DL-ML-SEM; Morin et al., 2022). We used Mplus 8.7 with the MLR estimator (Muthén & Muthén, 2019). As illustrated in

Fig. 2, some students were lost from the sample during the first semester (from T1 to T2; 8.8 %) and during the transition to the second semester (from T2 to T3; 4.6 %). At T2, some students were unable to attend class because they had a fever or cough or because the family had the student miss class that day for Covid-19 reasons. To handle these missing cases (and < 0.1 % missing data), we applied the full information likelihood (FIML) method to generate unbiased estimates that allowed us to analyze the full student sample (N = 1084; Enders, 2010). To evaluate model fit, we used four 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 < than 0.08 and 0.06; for CFI and TLI, adequate and excellent fit are reflected by values >0.90 and 0.95 (Hu & Bentler, 1999).

The measurement model included 18 indicators for four latent variables. In evaluating the measurement model, we evaluated the fit of both the constrained and unconstrained models. We did this to test for multilevel measurement invariance (i.e., metric invariance; Morin et al., 2022). To make this evaluation, we compared the fit of the unconstrained measurement model (the indicators were free to vary) versus the fit of the constrained model (the indicators were invariant across both level [L1,L2] and time [T1, T2, T3]). If the constrained measurement model shows little or no decrement in the goodness-of-fit statistics compared fit of the unconstrained measurement model, this result supports multilevel and multiwave measurement invariance (Marsh et al., 2011).

In the test of the hypothesized model, we added experimental condition as an uncentered L2 T1 predictor (experimental = 1, control = 0), gender as a grand mean-centered L1 covariate, and grade level and class size as grand mean-centered L2 covariates (see Fig. 1). To test the status of "L2 T2 autonomy-supportive peer climate" as a mediator in H2, H3, and H4, we tested for indirect effects in two ways: (a) within Mplus (using the "model indirect" command) and (b) with Preacher and Selig's (2012) 95 % confidence interval bootstrapping procedure (i.e., if the 95 % CI does not contain 0, mediation is confirmed).

Intervention Effect on the Five Measures. In a series of five DL-ML-SEM analyses (one for each measure), we tested whether the intervention (i.e., experimental condition) increased perceived autonomy-supportive teaching (i.e., manipulation check), increased perceived autonomy-supportive peer climate (i.e., hypothesized mediator), and increased defending- and decreased passive- and reinforcing-bystanding (i.e., the three outcomes). To create each latent variable, we used the six items from the LCQ for the two autonomy support measures and the four items from each BBS scale for the three bystanding measures. In each analysis, we used the same growth analysis to test for a T1 to T3 increase (or decrease) in the measure (slope: T1 = 0, T2 = 1, T3 = 2). We used linear parameters (0,1,2) because our interest was in the overall T1 to T3 change in scores rather than in the trajectory of that growth. As before, gender was a L1 covariate, whereas grade level and class size were L2 covariates. Across each analysis, we tested for a condition *x* time interaction to evidence that the L2 score increased or decreased more in the experimental condition than it did in the control condition.

In a set of preliminary analyses, we checked the efficacy of our random assignment to condition by testing for group differences at T1 (baseline). As indicated in Figs. 3b, c, 4a, b, and c, T1 scores on these five dependent measures did not differ significantly by experimental condition (experimental T1 score vs. control T1 score), all *ts* < 1.

Intervention Effect on Teachers' Instructional Behaviors. We tested whether the raters scored experimental group teachers higher than control group teachers in their in-class use of autonomy-supportive instructional behaviors and lower in controlling behaviors. For these two manipulation checks, the teacher was the unit-of-analysis, the statistical test was a two-group independent *t*-test, and Cohen's *d* statistic provided effect size information.

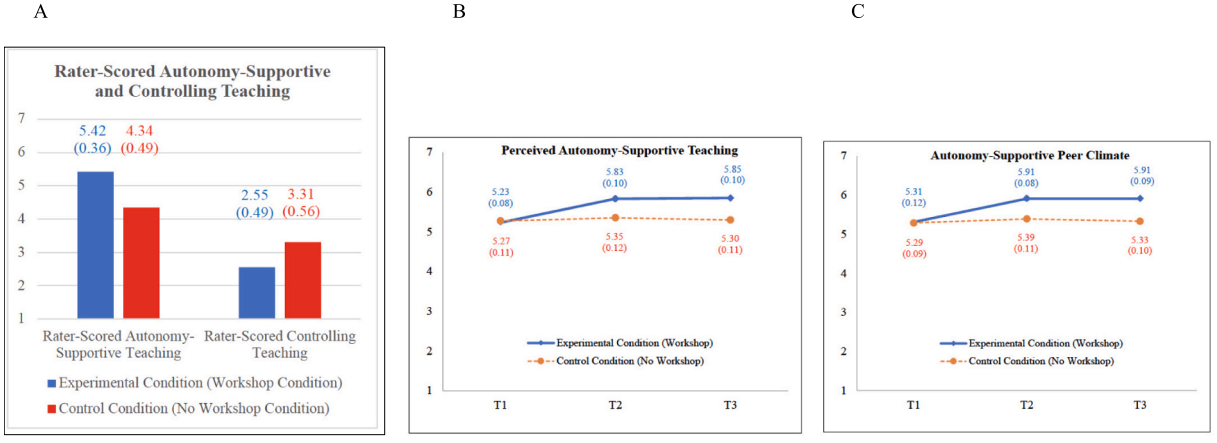

Fig. 3. Rater-Scored and Student Reported Autonomy-Supportive and Controlling Teaching by Experimental Condition (Study 1).
Note. The numbers in fig. A are classroom observers' objective scores (standard deviations in parentheses). The numbers in figs. B and C are student-reported class-average means (standard errors in parentheses). All effects control for student gender, grade level, and class size.

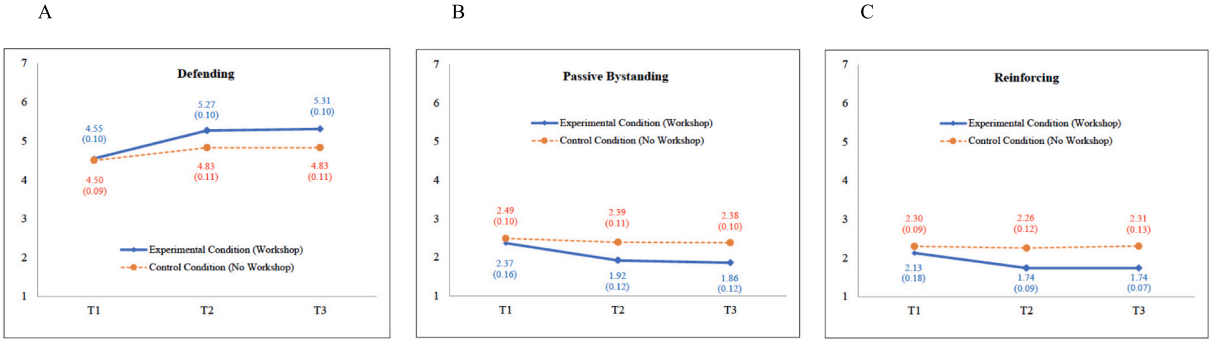

Fig. 4. Student Reported Bystanding Behaviors by Experimental Condition and Time of Assessment (Study 1).
Note. The numbers in the figures are student-reported class-average means (standard errors in parentheses). All effects control for student gender, grade level, and class size.

8. Results

8.1. Intervention effect on the two rater-scored instructional behaviors

The left side of Fig. 3A shows that raters scored experimental group teachers as engaging in more in-class autonomy-supportive instructional behaviors than did control group teachers ($M = 5.42$ vs. 4.34), $t(36) = 8.06$, $p < .001$, $d = 2.51$. The right side of Fig. 3a shows that raters scored experimental group teachers as engaging in less in-class controlling instructional behaviors than did control group teachers ($M = 2.55$ vs. 3.31), $t(36) = 4.73$, $p < .001$, $d = 1.44$.

8.2. Intervention effect on the five student-reported measures

Perceived Autonomy-Supportive Teaching. The DL-ML-SEM model with T1, T2, and T3 perceived autonomy-supportive teaching, experimental condition, and the three covariates fit the data reasonably well, $X^2(351) = 1049.08$, $p < .001$, RMSEA = 0.043, SRMR = 0.049, CFI = 0.934, and TLI = 0.929. The condition × time interaction (see Fig. 3b) was significant, $B = 0.65$, $SE = 0.12$, $t = 5.58$, $p < .001$ (estimated T1–T3 M Δ: $+0.62_{Exp}$ vs. $+0.03_{Con}$) with a large effect size, L2 $R^2 = 0.44$. That is, perceived autonomy-supportive teaching increased more from T1 to T3 for students in the experimental condition (the baseline score of 5.23 increased +0.62 to 5.85 by T3) than it did for students in the control condition (5.27 increased only +0.03 to 5.30).

Perceived Autonomy-Supportive Peer Climate. A parallel DL-ML-SEM model applied to the autonomy-supportive peer climate measure also fit the data reasonably well, $X^2(351) = 1020.50$, $p < .001$, RMSEA = 0.043, SRMR = 0.058, CFI = 0.954, and TLI = 0.950. The condition × time interaction (Fig. 3c) was significant, $B = 0.60$, $SE = 0.14$, $t = 4.30$, $p < .001$ (estimated T1-T3 M Δ: $+0.60_{Exp}$ vs. $+0.04_{Con}$) with a large effect size, L2 $R^2 = 0.36$. That is, the perceived autonomy-supportive peer climate increased more from T1 to T3 for students in the experimental condition (5.31 increased +0.60 to 5.91) than it did for students in the control condition (5.29 increased +0.04 to 5.33).

Defending. The DL-ML-SEM model for defending fit the data reasonably well, $X^2(157) = 471.58$, $p < .001$, RMSEA = 0.043, SRMR = 0.051, CFI = 0.963, and TLI = 0.958. The condition × time interaction (Fig. 4a) was significant, $B = 0.56$, $SE = 0.16$, $t = 3.51$, $p < .001$ (estimated T1-T3 M Δ: $+0.76_{Exp}$ vs. $+0.33_{Con}$) with a large effect size, L2 $R^2 = 0.34$. That is, defending increased more from T1 to T3 for students in the experimental condition (4.55 increased +0.76 to 5.31) than it did for students in the control condition (4.50 increased +0.33 to 4.83).

Passive Bystanders. The DL-ML-SEM model for passive bystanders fit the data well, $X^2(157) = 314.39$, $p < .001$, RMSEA = 0.030, SRMR = 0.050, CFI = 0.981, and TLI = 0.978. The condition × time interaction (Fig. 4b) was significant, $B = -0.45$, $SE = 0.19$, $t = 2.26$, $p = .024$ (estimated T1-T3 M Δ: -0.51_{Exp} vs. -0.11_{Con}) with a moderate effect size, L2 $R^2 = 0.23$. That is, passive-bystanding decreased more from T1 to T3 for students in the experimental condition (2.37 decreased −0.51 to 1.86) than it did for students in the control condition (2.49 decreased.

−0.11 to 2.38).

Reinforcing. The DL-ML-SEM model for reinforcing fit the data reasonably well, $X^2(157) = 480.96$, $p < .001$, RMSEA = 0.044, SRMR = 0.050, CFI = 0.960, and TLI = 0.954. The condition × time interaction (Fig. 4c) was significant, $B = -0.37$, $SE = 0.15$, $t = 2.41$, $p = .016$ (estimated T1-T3 M Δ: -0.39_{Exp} vs. $+0.01_{Con}$) with a moderate effect size, L2 $R^2 = 0.28$. That is, reinforcing decreased more from T1 to T3 for students in the experimental condition (2.13 decreased −0.39 to 1.74) than it did for students in the control condition (2.30 increased +0.01 to 2.31).

8.3. Hypothesized model test

Measurement Invariance. The unconstrained measurement model fit the data well, $X^2(1114) = 2068.61$, $p < .001$, RMSEA = 0.028, SRMR = 0.025, CFI = 0.972, and TLI = 0.968. The measurement model that constrained the indicators to be invariant across

both level (L1,L2) and time (T1, T2, T3) also fit the data well, and it did so with no decrement in the fit indices, $X^2(1156) = 2092.54$, $p < .001$, RMSEA = 0.027, SRMR = 0.025, CFI = 0.973, TLI = 0.970. These analyses establish measurement invariance for the measurement model.

Hypothesized Model. The hypothesized model fit the data well, $\chi^2(1302) = 2526.84$, $p < .001$, RMSEA = 0.029, SRMR = 0.041, CFI = 0.966, and TLI = 0.964. Fig. 5 shows the unstandardized beta weights (with standard errors) for the hypothesized paths, autoregressive effects, and statistical controls. Table 1 shows the L1 and L2 correlations for the latent variables and statistical controls included in the hypothesized model.

L2 Classroom-Level Effects. Consistent with H1, the intervention effect on T2 autonomy-supportive peer climate was positive and significant ($B = 0.55$, $SE = 0.11$, $t = 4.97$, $p < .001$), controlling for T1 autonomy-supportive climate and the set of statistical controls (see Fig. 5). This means that teachers in the experimental condition created a more autonomy-supportive climate than did teachers in the control condition.

Consistent with H2, the more autonomy-supportive students perceived the T2 climate to be, the more was their T3 defending ($B = 0.99$, $SE = 0.19$, $t = 5.32$, $p < .001$), controlling for T1 defending, T1 climate, experimental condition, and the statistical controls. Furthermore, this T2 classroom climate effect mediated the otherwise direct effect of the experimental condition on T3 defending (as per Fig. 4a): Indirect mediated effect = 0.55, $SE = 0.16$, $t = 3.32$, $p = .001$; and the 95 % CI from bootstrapping did not include 0 (0.28, 0.87).

Consistent with H3, the more autonomy-supportive students perceived the T2 peer climate to be, the less was their T3 passive-bystanding ($B = -0.95$, $SE = 0.16$, $t = 6.08$, $p < .001$), controlling for T1 passive-bystanding, T1 climate, experimental condition, and the statistical controls. Furthermore, this T2 classroom climate effect mediated the otherwise direct effect of the experimental condition on T3 passive-bystanding (as per Fig. 4b): Indirect mediated effect = −0.52, $SE = 0.12$, $t = 4.28$, $p < .001$; and the 95 % CI from bootstrapping did not include 0 (−0.28, −0.81).

Consistent with H4, the more autonomy-supportive students perceived the T2 peer climate to be, the less was their T3 reinforcing ($B = -0.73$, $SE = 0.20$, $t = 3.63$, $p < .001$), controlling for T1 reinforcing, T1 climate, experimental condition, and the statistical controls. Furthermore, this T2 classroom climate effect mediated the otherwise direct effect of the experimental condition on T3 reinforcing (as per Fig. 4c): Indirect mediated effect = −0.40, $SE = 0.14$, $t = 2.74$, $p = .006$; and the 95 % CI from bootstrapping did not include 0 (−0.16, −0.70).

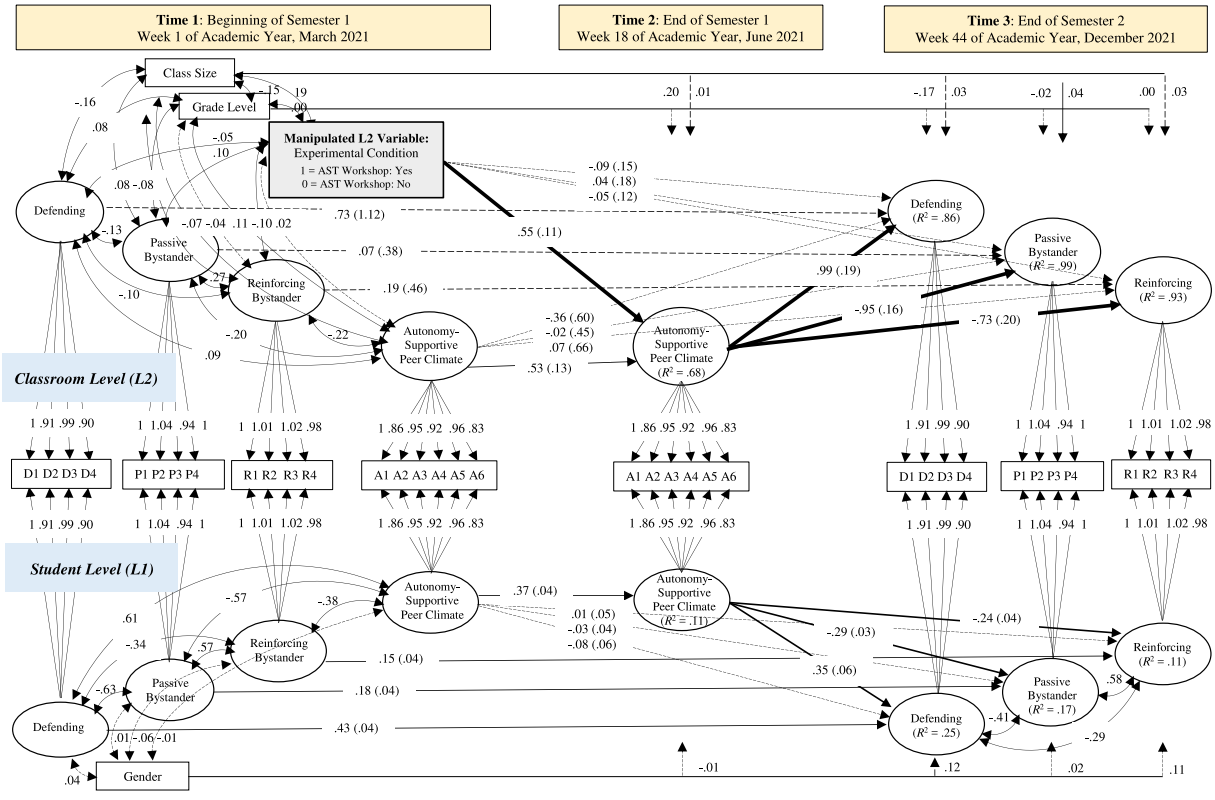

Fig. 5. Unstandardized Beta Weights (with Standard Errors) from the Test of the DL-ML-SEM Hypothesized Model (Study 1).
Note. D = Defending item; P = Passive item; R = Reinforcing item; A = Autonomy-Support Peer Climate item. Thick lines are hypothesized paths. Thin lines are statistical controls, autoregressive effects, and indicators of a latent construct. Rectangles denote observed variables; ovals represent latent variables. Overall Model Fit: $\chi^2(1302) = 2526.84$, $p < .001$, RMSEA = 0.029, SRMR = 0.041, CFI = 0.966, and TLI = 0.964.

Table 1
Correlations and Descriptive Statistics for the L1 and L2 Latent Variables and Statistical Controls (Study 1).

L1: Student-Level

	Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.
Time 1 Baseline										
1.	Autonomy-Supportive Peer Climate	–	0.47	−0.49	−0.42	0.33	0.22	−0.22	−0.14	−0.12
2.	Defending		–	−0.46	−0.31	0.16	0.40	−0.15	−0.08	−0.03
3.	Passive-Bystander			–	0.59	−0.14	−0.20	0.25	0.12	0.13
4.	Reinforcing				–	−0.16	−0.14	0.16	0.17	0.20
Time 2 Mediator										
5.	Autonomy-Supportive Peer Climate					–	0.34	−0.36	−0.30	−0.06
Time 3 Outcomes										
6.	Defending						–	−0.43	−0.33	0.00
7.	Passive-Bystander							–	0.66	0.09
8.	Reinforcing								–	0.14
Statistical Controls										
9.	Gender									–
Descriptive Statistics										
	Mean	5.28	4.50	2.44	1.97	5.62	5.07	2.12	1.87	0.55
	Standard Deviation	1.08	1.27	1.24	1.09	1.20	1.43	1.23	1.11	0.55
	Skewness	−0.35	−0.50	1.18	1.50	−0.74	−0.58	1.12	1.46	−0.19
	Kurtosis	−0.31	−0.29	0.93	1.92	0.14	−0.30	0.65	1.84	−1.96

Note. $N = 1084$ students. Correlations $r > 0.06$ are statistically significant ($p < .05$).

L2: Classroom-Level

	Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
					Time 1 Baseline							
1.	Experimental Condition	–	0.02	0.00	−0.05	0.10	0.62	0.46	−0.50	−0.53	0.00	0.19
2.	Autonomy-Supportive Peer Climate		–	0.75	−0.93	−0.92	0.49	0.49	−0.58	−0.59	0.22	0.14
3.	Defending			–	−0.83	−0.50	0.36	0.57	−0.44	−0.41	0.19	−0.40
4.	Passive-Bystander				–	0.92	−0.46	−0.52	0.55	0.56	−0.09	0.08
5.	Reinforcing					–	−0.45	−0.40	0.54	0.58	−0.03	−0.10
					Time 2 Mediator							
6.	Autonomy-Supportive Peer Climate						–	0.83	−0.94	−0.90	0.20	0.19
					Time 3 Outcomes							
7.	Defending							–	−0.87	−0.75	0.04	−0.04
8.	Passive-Bystander								–	0.96	−0.26	0.11
9.	Reinforcing									–	−0.20	0.10
					Statistical Controls							
10.	Grade Level										–	−0.15
11.	Class Size											–
					Descriptive Statistics							
	Mean	0.50	5.28	4.52	2.43	1.98	5.60	5.04	2.16	1.90	0.34	27.1
	Standard Deviation	0.50	0.46	0.39	0.57	0.63	0.50	0.53	0.52	0.49	0.48	3.2
	Skewness	0.07	−0.60	0.46	0.86	1.61	−0.27	0.13	0.54	1.05	0.69	−0.19
	Kurtosis	−2.08	−0.02	−0.71	0.64	2.18	0.14	−0.36	−0.19	0.86	−1.60	−0.56

Note. $k = 38$ teachers. Correlations $r > 0.34$ are statistically significant ($p < .05$).

L1 Student-Level Effects. As shown in the lower half of Fig. 5, at the L1 level, the more autonomy-supportive individual students' perceived the T2 peer climate to be, the more was their personal T3 defending ($B = 0.35$, $SE = 0.06$, $t = 5.72$, $p < .001$) and the less was their personal T3 passive ($B = -0.29$, $SE = 0.03$, $t = 8.64$, $p < .001$) and T3 reinforcing ($B = -0.24$, $SE = 0.04$, $t = 5.87$, $p < .001$) bystanding, controlling for T1 climate, T1 baseline score for the corresponding bystander behavior, and the statistical controls.

9. Discussion

Study 1 produced three new findings. First, professionally trained autonomy-supportive teachers were able to create a highly supportive peer-to-peer classroom climate ($M = 5.91$ on a 1–7 scale; see Fig. 3c). Second, the students in these supportive climates engaged in more defending and in less passive and reinforcing bystanding (see Fig. 5). Third, the mediation analysis showed that the reason why students showed more defending and less passive and reinforcing bystanding in the classes of professionally trained autonomy-supportive teachers was because these teachers were able to first create a more autonomy-supportive classroom climate.

In Study 1, teachers who participated in the AST workshop learned both greater autonomy support and lesser teacher control. Recognizing this, Study 2 sought to include measures of both autonomy-supportive and controlling teaching (not just autonomy-supportive teaching, as in Study 1). Thus, Study 2 featured two predictive variables in the hypothesized model. Similarly, Study 1 assessed only how supportive students perceived the classroom climate to be. So, in Study 2, we assessed both how supportive and how

conflictual the peer-to-peer classroom climate was. Thus, Study 2 featured two predictors and two explanatory mediators.

10. Study 2

The hypothesized model for Study 2 is illustrated in Fig. 6. The top half of the model corresponds closely to the hypothesized model tested in Study 1 (i.e., autonomy-supportive teaching ➔ supportive climate ➔ bystanding). The bottom half of the model adds potentially new explanatory processes (i.e., controlling teaching ➔ conflictual climate ➔ bystanding). We added controlling teaching because we suspected that although autonomy-supportive teaching facilitates supportive norms and defending bystanding, it might be that controlling teaching exacerbates conflictual norms and passive and reinforcing bystanding.

As shown on the left side of Fig. 6, we hypothesized that students' perception of their teacher as autonomy supportive would predict teachers' rating of a supportive climate (Hypothesis 1) whereas students' perception of their teacher as controlling would predict teachers' rating of a conflictual climate (Hypothesis 2). As shown on the right side of Fig. 6, we further hypothesized that teacher-rated supportive climate would explain relatively high defending- (Hypothesis 3), low passive- (Hypothesis 4), and low reinforcing- (Hypothesis 5) bystanding. Similarly, we hypothesized that teacher-rated conflictual climate would explain relatively low defending- (Hypothesis 6), high passive- (Hypothesis 7), and high reinforcing- (Hypothesis 8) bystanding. To connect the two student-reported predictor variables in the hypothesized model with the three teacher-rated outcomes, we conducted mediation analyses for H3 through H8, using both Mplus' "model indirect" command and the bootstrapping "95 % CI" procedure (as per Study 1). These mediation tests do not, however, suggest causal effects as we assessed the hypothesized mediators and outcomes within the same wave of data collection (T2) at the end of the course.

11. Method

11.1. Participants

Eleven PE teachers (four women and seven men) from 11 different schools (three middle and eight high-schools) in Seoul, Korea served as teacher-participants. Teachers were, on average, 34.3 years old ($SD = 5.9$, range = 27–42) with 6.0 years ($SD = 4.3$, range = 1–13) of teaching experience. To increase the number of classrooms (L2 units) included in the analysis, we collected data in multiple classrooms for each teacher. In the end, this data collection effort yielded data from two classrooms for seven teachers, one classroom for three teachers, and four classrooms for one teacher (i.e., 11 teachers, 25 classrooms). This data collection procedure allowed us to increase the number of classrooms (L2 units) included in the analyses, thereby increasing our statistical power. In these 25 classrooms

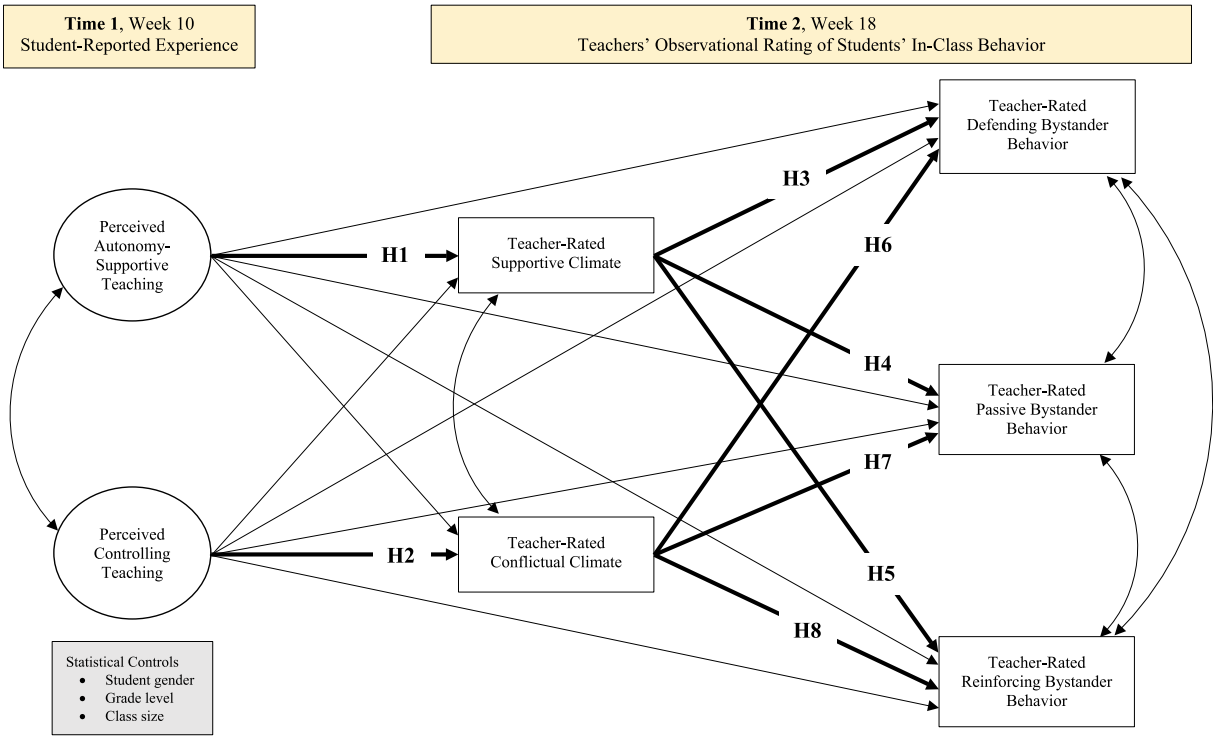

Fig. 6. Hypothesized Model (Study 2).
Note. Thick lines are hypothesized paths. Thin lines are statistical controls.

were 629 middle-class Korean adolescents (M_{age} = 15.0 years, SD = 1.6, range = 12–18), including 398 (63.3 %) females and 231 (36.7 %) males, and 233 (37.0 %) middle and 396 (63.0 %) high schoolers. The average class size was 25.5 students per class. None of the teachers and none of the students from Study 1 served as participants in Study 2.

11.2. Procedure and research design

The research protocol was approved by Korea University's ethics committee. The research design was a correlational study in which students completed the study questionnaire during Week 10 of an 18-week semester; teachers completed their ratings during the last week of the semester (Week 18). Because the hypothesized model in Study 2 was a mediation model, we acknowledge that a methodologically stronger research design would have involved asking teachers to make the classroom climate ratings (i.e., mediators) at a time well before they made the bystanding ratings (i.e., outcomes). However, we decided to ask teachers to provide only end-of-semester ratings so to decrease teacher burden (and to increase teacher participation), a decision that we acknowledge limits our directional interpretation of the hypothesized mediation effects.

11.3. Measures

Teachers' Observational Ratings. We gave each teacher a two-page class roster that described the rating scale (1 = *strongly disagree*, 7 = *strongly agree*) and listed five columns of student behaviors to rate. Each class roster included 30 rows. In each row, the rating sheet listed the student ID number for each student in that teacher's PE class. Page 1 featured the two classroom climate ratings whereas page 2 featured the three bystander ratings. This teacher rating sheet was in Korean, but an English-translated version appears on the OSF project site.

Classroom Climate Ratings. Page 1 described the situational context ("When interacting with his/her PE classmates, this student…"), two student behaviors for teachers to rate (i.e., supportive and conflictual), and exemplary behaviors to operationally define supportive (i.e., understands, accepts, and shows caring) and conflictual (i.e., rejects, blames, and belittles). We selected the three supportive exemplars from the highest loading (i.e., most representative) items on the Tripartite Measure of Interpersonal Behavior's (TMIB; Bhavsar et al., 2019) eight-item Need Support scale. We selected the three conflictual exemplars from the highest loading items on the TMIB's eight-item Need Thwart scale. As shown by Bhavsar et al. (2019), TMIB Need Support scores predict student perceptions of the extent to which a relationship partner actively supports their needs and concerns, and TMIB Need Thwart scores predict student perceptions of the extent to which a relationship partner actively thwarts their needs and concerns. In the present study, teachers, on average, rated their students higher on supportive (M = 5.17, SD = 1.39) than on conflictual (M = 1.77, SD = 0.97).

Bystanding Ratings. Page 2 described the situational context ("When a PE classmate is being bullied by other students, this student…"), three student behaviors for teachers to rate (i.e., defending, passive, and reinforcing), and exemplary behaviors to operationally define defending (i.e., defends, does something to stop the bullying, and tells a teacher), passive (i.e., walks away from, ignores, and just stands by and minds their own business), and reinforcing (i.e., laughs, cheers, and encourages the one bullying) bystanding. We selected these representative behaviors from a review of all existing measures of bystanding behavior in educational settings to identify those that best represented the conceptual definitions of bystanding behavior reviewed in the Introduction (see paragraph 2). In the present study, teachers, on average, rated their students higher on defending (M = 5.35, SD = 1.19) than on passive (M = 2.19, SD = 1.18) and reinforcing (M = 1.46, SD = 0.87).

Students' Self-Reports. The student questionnaire assessed perceived autonomy-supportive teaching and perceived controlling teaching (i.e., the two predictors in the hypothesized model; see Fig. 6). For perceived autonomy-supportive teaching, students completed the three-item Autonomy Support scale from the TMIB (Bhavsar et al., 2019; α = 0.88; e.g., "My PE teacher shows that he/she understands my perspective"; ω = 0.78 in the present study). For perceived controlling teaching, students completed the three-item Autonomy Thwart scale from the TMIB (Bhavsar et al., 2019; α = 0.77; e.g., "My PE teacher tries to control everything I do"; ω = 0.78 in the present study). The TMIB Autonomy Support scores have been shown to predict students' experiences of autonomy need satisfaction, whereas TMIB Autonomy Thwart scores have been shown to predict students' experiences of autonomy need frustration (r = 0.50, $p < .001$ and r = 0.59, $p < .001$, respectively; Bhavsar et al., 2019).

11.4. Data analyses

The data had a two-level structure with students' data (Level 1, N = 629) nested within classrooms (Level 2, k = 25) and nested further within teachers (a cross-classified Level 2, k = 11). Given this data structure, we used a multilevel structural equation modeling analysis to test the hypothesized model. We used Mplus 8.7 (Muthén & Muthén, 2019), the "model = complex" command to handle the nested structure of the data, the maximum likelihood-robust estimator (MLR), and the FIML estimation procedure for handling missing data, which were rare (< 0.1 %). To evaluate model fit, we again used the RMSEA, SRMR, CFI, and TLI goodness-of-fit statistics. In the hypothesized model, we entered the two student-reported scores as latent variables and entered the five teacher ratings as single-item manifest or observed variables (Fig. 6). We entered student gender (female = 1, male = 0), grade level (high = 1, middle = 0), and class size (M = 25.6) as statistical controls.

12. Results

Table 2 reports the descriptive statistics and intercorrelations for the variables included in the hypothesized model. The

hypothesized model fit the data very well, $\chi^2(46) = 56.89$, $p = .130$, RMSEA = 0.019, SRMR = 0.031, CFI = 0.986, and TLI = 0.974. Fig. 7 shows the unstandardized beta weights (with standard errors) for the hypothesized paths and statistical controls.

12.1. Motivating style effects on supportive-conflictual

Consistent with H1, student-reported perceived autonomy-supportive teaching significantly predicted teacher-rated supportive climate ($B = 0.23$, $SE = 0.11$, $t = 2.10$, $p = .036$), controlling for perceived controlling teaching ($B = -0.13$, $p = .181$), gender ($B = 0.10$, $p = .414$), grade level ($B = 0.30$, $p = .437$), and class size ($B = 0.01$, $p = .821$). Consistent with H2, student-reported perceived controlling teaching significantly predicted teacher-rated conflictual climate ($B = 0.11$, $SE = 0.05$, $t = 2.04$, $p = .041$), controlling for perceived autonomy-supportive teaching ($B = -0.10$, $p = .091$), gender ($B = 0.32$, $p = .011$), grade level ($B = 0.12$, $p = .318$), and class size ($B = -0.02$, $p = .272$; see Fig. 7).

12.2. Supportive-conflictual effects on bystanding behaviors

For teacher-rated defending, supportive climate was an individually significant positive predictor ($B = 0.28$, $SE = 0.10$, $t = 2.75$, $p = .006$) whereas conflictual climate was an individually significant negative predictor ($B = -0.22$, $SE = 0.08$, $t = 2.65$, $p = .008$), controlling for perceived autonomy-supportive teaching ($B = 0.05$, $p = .380$), perceived controlling teaching ($B = -0.02$, $p = .688$), gender ($B = -0.02$, $p = .878$), grade level ($B = -0.34$, $p = .138$), and class size ($B = -0.04$, $p = .242$). In the test for mediation, supportive climate and conflictual climate together only marginally mediated the direct effect of autonomy-supportive teaching on defending: Total indirect effect = 0.09, $SE = 0.04$, $t = 1.95$, $p = .052$. However, in the bootstrapping analysis, the 95 % CI did not include 0 for either supportive climate (0.002, 0.156, via autonomy-supportive teaching) or conflictual climate (−0.059, −0.0002, via controlling teaching), thereby suggesting double mediation (see Fig. 7).

For teacher-rated passive bystanding, conflictual climate was an individually significant positive predictor ($B = 0.26$, $SE = 0.08$, $t = 3.21$, $p = .001$) whereas supportive climate was an individually significant negative predictor ($B = -0.29$, $SE = 0.08$, $t = 3.71$, $p < .001$), controlling for perceived autonomy-supportive teaching ($B = -0.04$, $p = .511$), perceived controlling teaching ($B = 0.02$, $p = .771$), gender ($B = 0.08$, $p = .450$), grade level ($B = 0.90$, $p < .001$), and class size ($B = 0.12$, $p < .001$). In the test for mediation, supportive climate and conflictual climate together only marginally mediated the direct effect of controlling teaching on passive-bystanding: Total indirect effect = 0.06, $SE = 0.03$, $t = 1.84$, $p = .065$. However, in the bootstrapping analysis, the 95 % CI did not include 0 for either supportive climate (−0.147, −0.003 via autonomy-supportive teaching) or conflictual climate (0.002, 0.066 via controlling teaching), thereby suggesting double mediation.

For teacher-rated reinforcing, conflictual climate was an individually significant positive predictor ($B = 0.37$, $SE = 0.06$, $t = 6.13$, $p < .001$) whereas supportive climate was not ($B = -0.01$, $SE = 0.02$, $t = 0.37$, $p = .710$), controlling for perceived autonomy-supportive teaching ($B = 0.06$, $p = .162$), perceived controlling teaching ($B = -0.02$, $p = .965$), gender ($B = 0.27$, $p = .004$), grade level ($B = 0.03$, $p = .809$), and class size ($B = -0.02$, $p = .194$). In the test for mediation, supportive climate and conflictual climate together only marginally mediated the direct effect of controlling teaching on reinforcing: Total indirect effect = 0.04, $SE = 0.02$, $t = 1.81$, $p = .070$. In the bootstrapping analysis, the 95 % CI for conflictual climate (via controlling teaching) did not include 0 (0.002, 0.084), thereby suggesting single mediation, whereas the 95 % CI for supportive climate did include 0 (−0.014, 0.009).

Table 2
Correlations and Descriptive Statistics for the Variables and Statistical Controls Included in the Hypothesized Model (Study 2).

	Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Variables in Hypothesized Model											
1.	Student's Perceived Autonomy-Supportive Teaching	–	−0.32	0.18	−0.13	0.13	−0.13	0.02	0.02	−0.10	0.16
2.	Student's Perceived Controlling Teaching		–	−0.13	0.13	−0.10	0.10	0.02	0.01	0.05	−0.01
3.	Teacher-Rated Supportive Climate			–	−0.29	0.37	−0.37	−0.12	0.02	0.09	0.01
4.	Teacher-Rated Conflictual Climate				–	−0.28	0.32	0.44	0.16	0.05	−0.06
5.	Teacher-Rated Defending Bystander Behavior					–	−0.50	−0.09	−0.01	−0.10	−0.07
6.	Teacher-Rated Passive Bystander Behavior						–	0.11	0.01	0.29	0.27
7.	Teacher-Rated Reinforcing Bystander Behavior							–	0.21	0.03	−0.11
Statistical Controls											
8.	Gender								–	−0.11	−0.15
9.	Grade Level									–	−0.16
10.	Class Size										–
Descriptive Statistics											
	Mean	5.32	1.95	5.17	1.77	5.35	2.19	1.46	0.37	0.63	25.6
	Standard Deviation	1.04	1.00	1.39	0.97	1.19	1.18	0.87	0.48	0.48	3.5
	Skewness	−0.41	1.11	−1.03	1.48	−0.60	1.03	2.30	0.55	−0.54	0.05
	Kurtosis	0.55	1.16	1.06	2.38	0.36	1.05	5.36	−1.70	−1.71	−1.28

Note. $N = 629$ students. Correlations $r > 0.09$ are statistically significant ($p < .05$).

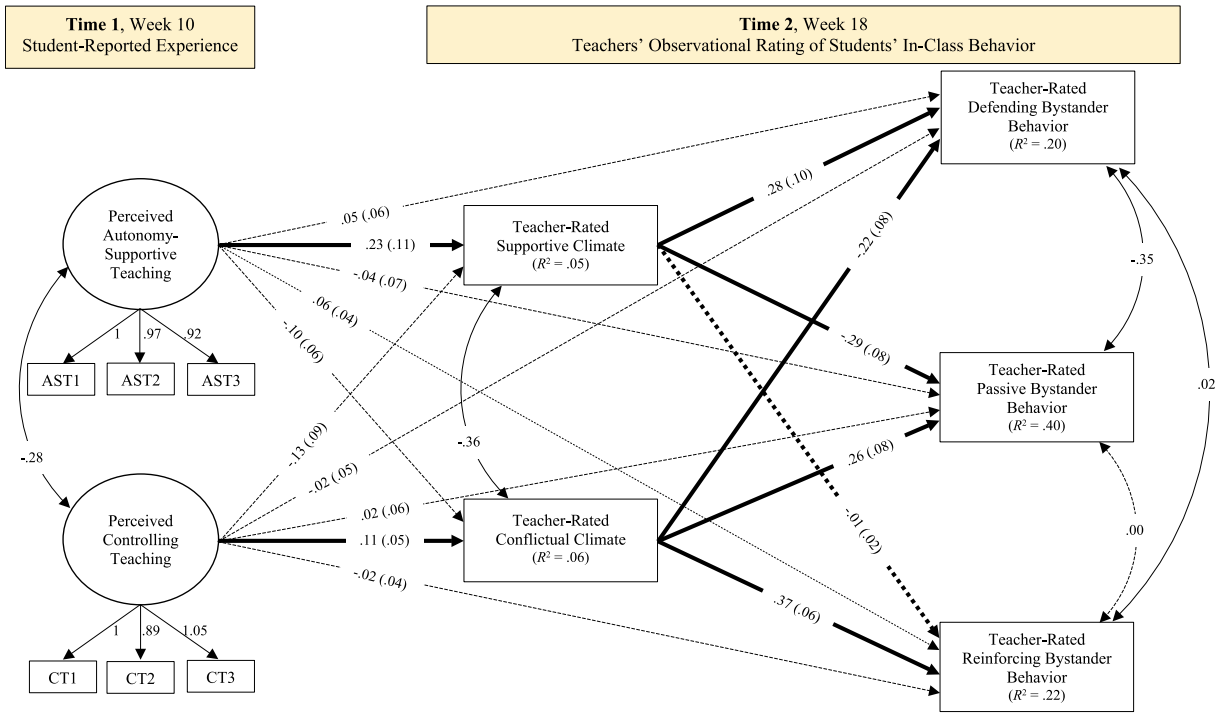

Fig. 7. Unstandardized Beta Weights (with Standard Errors) from the Test of the Hypothesized Model (Study 2).
Note. AST1 = autonomy-supportive teaching indicator #1; CT1 = controlling teaching indicator #1; R^2 = Proportion of explained variance. Thick lines are hypothesized paths. Thin lines are statistical controls or indicators of a latent construct. Solid lines represent statistically significant effects, $p < .05$, whereas dashed lines represent statistically non-significant paths. Effects of statistical controls (i.e., gender, grade level, and class size) are not shown in the figure, but each effect is reported in the text. Overall Model Fit: $\chi^2(46) = 56.89$, $p = .130$, RMSEA = 0.019, SRMR = 0.031, CFI = 0.986, and TLI = 0.974.

13. Discussion

Study 2 expanded on the Study 1 findings. Autonomy-supportive teaching once again predicted a supportive climate, which again predicted relatively high defending (and relatively low passive-bystanding) but not low reinforcing (i.e., compare the top half of Fig. 5 with the top half of Fig. 7). As in Study 1, the Study 2 mediation analyses confirmed that it was the supportive climate, rather than the teacher per se, that explained the bystanding behavior. One way that Study 2 added to the Study 1 findings was that it included perceived controlling teaching as a second predictor variable. It was perceived controlling teaching, rather than perceived autonomy-supportive teaching, that best predicted a conflictual climate. Another way that Study 2 added to the Study 1 findings was that it included a measure of the conflictual climate. In the Study 2 findings, supportive climate and conflictual climate were two independent and roughly equal predictors of defending and passive bystanding, but conflictual climate was the only individually significant predictor of reinforcing-bystanding.

14. General discussion

We sought to identify the antecedents of peer bystanding and to understand why students, over time, tend toward or away from each bystanding behavior. To do so, we integrated a social-ecological perspective into the self-determination theory framework to show how social forces and group dynamics (i.e., classroom climate) direct students toward or away from defending, passive, and reinforcing bystanding. The findings from both studies highlighted the predictive role of the prevailing classroom climate that emerged out of the teacher's motivating style. In Study 1, we used a cluster randomized control research design to show that this teacher-to-climate relationship was causal. Manipulated motivating style effected the prevailing classroom climate. Specifically, when teachers were professionally trained to be more autonomy supportive and less controlling, then the social contextual climate that formed in their classroom became increasingly supportive. Once established, this supportive climate then tended students toward greater defending, lesser passive-bystanding, and lesser reinforcing.

In Study 2, we expanded this focus on autonomy-supportive teaching and a supportive climate to include a parallel and supplementary focus on controlling teaching and a conflictual climate. Although the methodology employed in Study 2 was not as rigorous as that employed in Study 1, Study 2 still produced some suggestive findings. Study 2 suggested that controlling (not just autonomy-supportive) teaching and a conflictual (not just a supportive) climate play additionally important roles. A conflictual climate was a particularly important predictor of reinforcing bystanding.

15. Autonomy-supportive teaching, supportive climate, and defending bystanders

Study 1 showed that greater autonomy-supportive teaching led to the emergence of a supportive peer-to-peer classroom climate, which then strongly affected bystanding behaviors. The classroom climate effect on bystanding was so strong as to reach a near one-to-one correspondence (see Fig. 5), as a mid-year supportive climate increased end-year defending and decreased end-year passive and reinforcing bystanding. To explain why these relations were so strong, we suggest that defending is a tell-tale sign, symptom, or out-growth of a supportive peer climate, whereas passive-bystanding and, especially, reinforcing-bystanding are tell-tale signs, symptoms, or out-growths of a conflictual peer climate. We further suggest that future research may show that these two social processes are so "joined at the hip" that future longitudinal research may show that classroom climate and bystanding behavior could be reciprocally related and mutually reinforcing (e.g., Laninga-Wijnen et al., 2023).

It is worth noting that passive-bystanding was not neutral bystanding. Instead, passive-bystanding correlated strongly positively with reinforcing-bystanding and strongly negatively with defending-bystanding (see Table 1). This suggests that ignoring the situation or walking away is not neutral bystanding. Presumably, those participating in the bullying take another's passive lack of involvement as silent approval for their actions. In the social psychological literature on helping behavior, ignoring aggression and allowing it to proceed unchecked against a victim is called the bystander-apathy effect (Fischer et al., 2011). Thus, the defending role was the only constructive bystander behavior to emerge in the present study.

Several recent empirical investigations have been conducted to better understand the interrelationships among autonomy-supportive teaching, classroom climate, bystanding, and bullying-victimization. In the first of these studies, researchers conducted a RCT to show that students' in-class antisocial behavior emerged mostly out of a problematic social ecology (i.e., conflictual classroom climate) rather than out of problematic personal motivations (i.e., need frustration; Cheon et al., 2022). A follow-up RCT extended this finding to show that in-class prosocial behavior emerged mostly out of a favorable social ecology (i.e., supportive classroom climate) rather than out of favorable personal motivations (i.e., need satisfaction; Cheon, Reeve, Marsh, & Jang, 2023). A third RCT expanded this focus on episodic antisocial behavior to include recurring bullying-victimization; it showed that an intervention-enabled rise in a supportive classroom climate significantly decreased semester-long victimization (Cheon, Reeve, & Marsh, 2023). Still, the question remained as to whether defending decreased bullying-victimization. This relationship was clarified by a fourth longitudinal study that showed that defending predicted relatively low victimization in supportive classroom climates while it predicted relatively high victimization in conflictual classroom climates (i.e., a moderating effect; Malamut et al., 2023).

Collectively, what these studies seem to show is that one viable route to decrease bullying-victimization is to first improve the quality of the prevailing classroom climate to the point that supporting and defending become normative. The present Study 1 showed that manipulated autonomy-supportive teaching could indeed improve the quality of the classroom climate (i.e., more supportive) and, in doing so, exert a facilitative effect on defending-bystanding. The present Study 2 added that future efforts to reduce reinforcing-bystanding will likely require not only a more supportive climate but also a less conflictual classroom climate. Fortunately, the Study 1 findings showed that teachers who participate in the AST workshop learned how to become both more autonomy supportive and less controlling (see Fig. 3a). Because of this finding, and when it is combined with the earlier-reviewed set of studies, there is now a need for an intervention study in which teachers first participate in an AST workshop while researchers assess perceived autonomy-supportive and controlling teaching and perceived supportive and conflictual classroom climate to explain the longitudinal rise and fall of both bystanding behavior and bullying-victimization.

16. Practical recommendations

Our findings affirm that educators can do something constructive about students' bystanding behavior—namely introduce an autonomy-supportive teacher. Autonomy-supportive teaching was shown to be a malleable aspect of teaching. This malleable aspect of teaching was able to generate a more supportive peer social ecology that, once established, defending bystanding tended to follow, as shown on the right side of Figs. 5 and 7. It is another important practical point to note that teachers who participated in the autonomy-supportive teaching workshop learned not only greater autonomy-supportive teaching but also (a) lesser controlling teaching and (b) how to transform their existing controlling instructional behaviors into autonomy-supportive replacements (Fig. 3a; see Reeve et al., 2022, for a review). This means that professionally trained teachers may be able to facilitate both aspects of the classroom climate (i.e., more supportive and less conflictual) that function as important social processes to escalate and maintain defending bystanding and to de-escalate and minimize passive and reinforcing bystanding. We emphasize, however, that such a longitudinal change in bystander behavior is not a direct teacher effect. Instead, the Study 1 findings showed that the teacher effect (i.e., experimental condition) was to change the prevailing T2 classroom climate. It was this change in the T2 classroom climate that fully explained the longitudinal change in T3 bystander behavior (i.e., a fully mediated effect).

To reduce bullying, it is tempting (even commonsensical) to take a direct approach—to suppress, punish, or try to change bullying behavior. Fifty years of research shows that this direct approach has a poor track record (Juvonen & Graham, 2014). This is partly because students who bully tend to react to a direct approach with reactance and defiance (Cunningham et al., 2016). We suggest another reason is because bullying tends to arise out of social processes (Espelage & Swearer, 2004). So, instead of (or perhaps in addition to) a direct approach to reduce bullying, we recommend the mobilization of social forces.

We suggest that teachers and school psychologists learn how to cultivate a caring and mutually supportive classroom climate at the beginning of the semester or academic year. These peer interactions and relationships tend students toward defending, but they further ensure that students' defending initiatives will have the support, endorsement, and backing of their classmates. When this is the case, the defending bystander feels support, rather than at risk. It is at that point, once a reliable, dependable, consensual, and consistently

supportive classroom climate prevails, that it makes sense to recommend the teacher, school counselor, school psychologist, or school at large then begin to work more directly with the individual students involved in bully-victim episodes (e.g., see Garandeau et al., 2023).

17. Limitations

We note four key limitations. First, our teachers and students were Koreans taking a PE course at the secondary-grade level. Additional research will be necessary to evaluate how robust and generalizable the findings from the present study are. For instance, our hypothesized model needs to be tested with samples from different countries, grade levels, and the full range of subject matters.

Second, after our Study 1 intervention, we did not conduct a social fidelity check to ask if teachers agreed with our recommended teaching practices or if they felt this way of teaching produced a positive change in their teaching effectiveness. However, we conducted manipulation checks to confirm that teachers used the workshop experience to teach in a highly autonomy supportive. We encourage future investigations to assess both post-intervention checks (i.e., social fidelity and experimental treatment effects).

Third, our focus on bystanding behavior leaves our investigation open to potential social desirability and causality concerns. As to social desirability, it is possible that students might be hesitant to self-report, for example, reinforcing-bystanding. Our Study 2 methodology somewhat offset this concern by asking teachers to rate students' in-class bystanding behaviors. As to causality, our RCT research design supports our causal claim that manipulated autonomy-supportive teaching increases a supportive classroom climate (in Study 1). But our longitudinal research design cannot support causal claims regarding the right side of Study 1's hypothesized model or regarding Study 2's hypothesized model. In the absence of a RCT design, or even with random assignment, alternative explanations of the observed effects are possible that may differ from implicit causal interpretations. Thus, for the non-manipulated longitudinal effects (e.g., climate ➔ bystanding behavior), we use the term "effect" in its generic sense.

Fourth, we focused narrowly on the classroom context. In contrast, many researchers recommend a whole-school approach (Ansary et al., 2015). These educators worry that the constructive teacher-student relationships and peer-group dynamics may not extend beyond the walls of one specific classroom when students leave to go to another class—or to the lunchroom, the playground, the sports field, or the school bus. In each new setting, the teacher-adult relationship changes, peers change, and the prevailing social norms may change. To address these "classroom vs. whole-school" concerns, additional research will be necessary.

18. Conclusion

With professional training, teachers can create a closely-knit and supportive peer-to-peer classroom climate. When students find themselves embedded within such a social ecology, they tend toward defending and away from passive and reinforcing bystanding.

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5A2A01076454). These funding sources had no other role other than financial support.

CRediT authorship contribution statement

Sung Hyeon Cheon: Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Johnmarshall Reeve:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Kyoung-Eun Yoo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Yong-Gwan Song:** Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization. **Herbert W. Marsh:** Writing – review & editing, Writing – original draft, Visualization, Software, Formal analysis, Conceptualization. **Hye-Ryen Jang:** Writing – review & editing, Visualization, Software, Resources, Formal analysis. **Youngsun Lee:** Validation, Supervision, Resources.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsp.2025.101431.

References

Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., & Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology, 111*(3), 497–521. https://doi.org/10.1037/edu0000293.

Ansary, N., Elias, M., Greene, M., & Green, S. (2015). Guidance for schools selecting antibullying approaches: Translating evidence-based strategies to contemporary implementation realities. *Educational Researcher, 44*(1), 27–36. https://doi.org/10.3102/0013189X14567534.

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.

Bhavsar, N., Ntoumanis, N., Quested, E., Gucciardi, D., Thogersen-Ntoumani, C., Ryan, R. M., … Bartholomew, K. (2019). Conceptualizing and testing a new tripartite measure of coach interpersonal behaviors. *Psychology of Sport and Exercise, 44*, 107–120. https://doi.org/10.1016/j.psychsport.2019.05.006.

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, 10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3.

Cheon, S. H., Reeve, J., Joo, W.-Y., Song, Y.-G., Ryan, R. M., & Jang, H. (2024). Two randomized control trials to help teachers develop physical education students' course-specific grit-perseverance and mental toughness. *Journal of Sport & Exercise Psychology, 46*(5), 266–282. https://doi.org/10.1123/jsep.2024.0102.

Cheon, S. H., Reeve, J., & Marsh, H. (2023). Autonomy-supportive teaching enhances prosocial and reduces antisocial behavior via classroom climate and psychological needs: A multilevel randomized control intervention. *Journal of Sport and Exercise Psychology, 45*(1), 26–40. https://doi.org/10.1123/jsep.2021-0337.

Cheon, S. H., Reeve, J., Marsh, H., & Jang, H.-R. (2023). Cluster randomized control trial to reduce peer victimization: An autonomy-supportive teaching intervention changes the classroom ethos to support defending bystanders. *American Psychologist, 78*(7), 856–872. https://doi.org/10.1037/amp0001130.

Cheon, S. H., Reeve, J., Marsh, H. W., & Song, Y.-G. (2022). Intervention-enabled autonomy-supportive teaching improves the PE classroom climate to reduce antisocial behavior. *Psychology of Sport and Exercise, 60*(1). Article 102174 https://doi.org/10.1016/j.psychsport.2022.102174.

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. https://doi.org/10.1016/j.psychsport.2017.11.010.

Cunningham, C. E., Mapp, C., Rimas, H., Cunningham, L., Mielko, S., Vaillancourt, T., & Marcus, M. (2016). What limits the effectiveness of antibullying programs? A thematic analysis of the perspective of students. *Psychology of Violence, 6*(4), 596–606. https://doi.org/10.1037/a0039984.

De Meyer, J., Soenens, B., Aelterman, N., De Bourdeaudhuij, I., & Haerens, L. (2016). The different faces of controlling teaching: Implications of a distinction between externally and internally controlling teaching for students' motivation in physical education. *Physical Education and Sport Pedagogy, 21*(6), 632–652. https://doi.org/10.1080/17408989.2015.1112777.

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology, 73*(5), 642–650. https://doi.org/10.1037/0022-0663.73.5.642.

Di Stasio, M. R., Savage, R., & Burgos, G. (2016). Social comparison, competition and teacher-student relationships in junior high school classrooms predicts bullying and victimization. *Journal of Adolescence, 53*, 206–207. https://doi.org/10.1016/j.adolescence.2016.10.002.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Espelage, D. L., & Swearer, S. M. (Eds.). (2004). *Bullying in American schools: A social-ecological perspective on prevention and intervention*. Lawrence Erlbaum Associates Publishers.

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*(4), 517–537. https://doi.org/10.1037/a0023304.

Flaspohler, P. D., Elfstrom, J. L., Vanderzee, K. L., Sink, H. E., & Birchmeier, Z. (2009). Stand by me: The effects of peer and teacher support in mitigating the impact of bullying on quality of life. *Psychology in the Schools, 46*(7), 636–649. https://doi.org/10.1002/pits.20404.

Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2021). What works in anti-bullying programs? Analysis of effective intervention components. *Journal of School Psychology, 85*(2), 37–85. https://doi.org/10.1016/j.jsp.2020.12/002.

Garandeau, C. F., Lee, I., & Salmivalli, C. (2014). Inequality matters: Classroom status hierarchy and adolescents' bullying. *Journal of Youth and Adolescence, 43*, 1123–1133.

Garandeau, C. F., Turunen, T., Saarento-Zaprudin, S., & Salmivalli, C. (2023). Effects of the KiVa anti-bullying program on defending behavior: Investigating individual-level mechanisms of change. *Journal of School Psychology, 99*, Article 101226. https://doi.org/10.1016/j.jsp.2023.101226.

Hawkins, D. L., Pepler, D. J., & Craig, W. M. (2001). Naturalistic observations of peer interventions in bullying. *Social Development, 10*, 512–527. https://doi.org/10.1111/1467-9507.00178.

Hinduja, S., & Patchin, J. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research, 14*(3), 206–221. https://doi.org/10.1080/13811118.2010.494133.

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*(4), 311–322. https://doi.org/10.1016/j.avb.2012.03.003.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. https://doi.org/10.1080/10705519909540118.

Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An introduction. *American Psychologist, 70*, 293–299. https://doi.org/10.1037/a0038928.

Jiménez-Barbero, J. A., Jiménez-Loaisa, A., González-Cutre, D., Beltrán-Carrillo, V. J., Llor-Zaragoza, L., & Ruiz-Hernández-Ruiz, J. A. (2020). Physical education and school bullying: A systematic review. *Physical Education and Sport Pedagogy, 25*, 79–100. https://doi.org/10.1080/17408989.2019.1688775.

Joussemet, M., & Grolnick, W. S. (2022). Parental consideration on children's experiences: A critical review of parenting constructs. *Journal of Family Theory and Review, 14*(4), 593–619, 10.111/jftr.12467.

Jungert, T., Piroddi, B., & Thornberg, R. (2016). Early adolescents' motivations to defend victims in school bullying and their perceptions of student-teacher relationships: A self-determination theory approach. *Journal of Adolescence, 53*(1), 75–90.

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.

Kärnä, 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.

Kärnä, A., Voeten, M., Poskiparta, E., & Salmivalli, C. (2010). Vulnerable children in varying classroom contexts: Bystanders' behaviors moderate the effects of risk factors on victimization. *Merrill-Palmer Quarterly, 56*, 261–282. https://digitalcommons.wayne.edu/mpq/vol56/iss3/4.

Kim, H., Han, Y., Song, J., & Song, T. M. (2019). Application of social big data to identify trends of school bullying forms in South Korea. *International Journal of Environmental Research and Public Health, 16*(14), 2596. https://doi.org/10.3390/ijerph16142596.

Lambe, L. J., Cioppa, V. D., Hong, I. K., & Craig, W. M. (2019). Standing up to bullying: A social ecological review of peer defending in offline and online contexts. *Aggression and Violent Behavior, 45*, 51–74. https://doi.org/10.1016/j.avb.2018.05.007.

Laninga-Wijnen, L., Malamut, S. T., Garandeau, C. F., & Salmivalli, C. (2023). Does defending affect adolescents' peer status, or vice versa? Testing the moderating effects of empathy, gender, and anti-bullying norms. *Journal of Research on Adolescence, 33*(3), 913–930. https://doi.org/10.1111/jora.12847.

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.

Malamut, S. T., Trach, J., Garandeau, C. F., & Salmivalli, C. (2023). Does defending victimized peers put youth at risk of being victimized? *Child Development, 94*(2), 380–394. https://doi.org/10.1111/cdev.13866.

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., Reeve, J., Guo, J., Pekrun, R., Parada, R. H., Parker, P. D., … Cheon, S. H. (2023). Overcoming limitations in peer-victimization research that impede successful intervention: Challenges and new directions. *Perspectives on Psychological Science, 18*(4), 812–828. https://doi.org/10.1177/17456916221112919.

Ministry of Education. (2021). *The results of the first survey on school violence in 2021*. Ministry of Education.

Morin, A. J. S., Blais, A.-R., & Chénard-Poirier, L. A. (2022). Doubly latent multilevel procedures for organizational assessment and prediction. *Journal of Business and Psychology, 37*(1), 47–72. https://doi.org/10.1007/s10869-021-09736-5.
Muthén, L., & Muthén, B. (2019). *Mplus user's guide*. Muthén & Muthén.
Nocentini, A., Menesini, E., & Salmivalli, C. (2013). Level and change of bullying behavior during high school: A multilevel growth curve analysis. *Journal of Adolescence, 36*, 495–505. https://doi.org/10.1016/j.adolescence.2013.02.004.
Ntoumanis, N., & Vazou, S. (2005). Peer motivational climate in youth sport: Measurement development and validation. *Journal of Sport and Exercise Psychology, 27*, 432–455.
Olweus, D. (1973). Personality and aggression. In J. K. Cole, & D. D. Jansen (Eds.), *Nebraska symposium on motivation* (pp. 260–321). University of Nebraska Press.
Patall, E. A., Dent, A. L., Oyer, M., & Wynn, S. R. (2013). Student autonomy and course value: The unique and cumulative roles of various teacher practices. *Motivation and Emotion, 37*(1), 14–32. https://doi.org/10.1007/s11031-012-9305-6.
Polanin, J. R., Espelage, D. L., & Pigott, T. D. (2012). A meta-analysis of school-based bullying prevention programs' effects on bystander intervention behavior. *School Psychology Review, 41*(1), 47–65. https://doi.org/10.1080/02796015.2012.12087375.
Poyhonen, V., Juvonen, J., & Salmivalli, C. (2010). What does it take to stand up for the victim of bullying? *Merrill-Palmer Quarterly, 56*, 143–163. https://doi.org/10.1353/mpq.0.0046.
Pozzoli, T., Gini, G., & Vieno, A. (2012). The role of individual correlates and class norms in defending and passive bystanding behavior in bullying: A multilevel analysis. *Child Development, 83*(6), 1917–1931.
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., Ryan, R. M., Cheon, S. H., Matos, L., & Kaplan, H. (2022). *Supporting students' motivation: Strategies for success*. Routledge. https://doi.org/10.4324/9781003091738.
Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.
Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15*, 112–120. https://doi.org/10.1016/j.avb.2009.08.007.
Salmivalli, C. (2014). Participant roles in bullying: How can peer bystanders be utilized in interventions? *Theory Into Practice, 53*(4), 286–292. https://doi.org/10.1080/00405841.2014.947222.
Salmivalli, C., Voeten, M., & Poskiparta, E. (2011). Bystanders matter: Associations between reinforcing, defending, and the frequency of bullying behavior in classrooms. *Journal of Clinical Child & Adolescent Psychology, 40*(5), 668–676. https://doi.org/10.1080/15374416.2011.597090.
Soenens, B., Sierens, E., Vansteenkiste, M., Dochy, F., & Goossens, L. (2012). Psychologically controlling teaching: Examining outcomes, antecedents, and mediators. *Journal of Educational Psychology, 104*(1), 108–120. https://doi.org/10.1037/a0025742.
Thornberg, R., Tenenbaum, L., Varjas, K., Meyers, J., Jungert, T., & Vanegas, G. (2012). Bystander motivation in bullying incidents: To intervene or not to intervene? *The Western Journal of Emergency Medicine, 13*(3), 247–252. https://doi.org/10.5811/westjem.2012.3.11792.
Van Ryzin, M. J., & Roseth, C. J. (2018). Cooperative learning in middle school: A means to improve peer relations and reduce victimization, bullying, and related outcomes. *Journal of Educational Psychology, 110*(8), 1192–1201. https://doi.org/10.1037/edu0000265.
Vansteenkiste, M., Aelterman, N., De Muynck, G.-J., Haerens, L., Patall, E., & Reeve, J. (2018). Fostering personal meaning and self-relevance: A self-determination theory perspective on internalization. *Journal of Experimental Education, 86*, 30–49.
Wu, W. C., Luu, S., & Luh, D. L. (2016). Defending behaviors, bullying roles, and their associations with mental health in junior high school students: A population-based study. *BMC Public Health, 16*, 1066. https://doi.org/10.1186/s12889-016-3721-6.