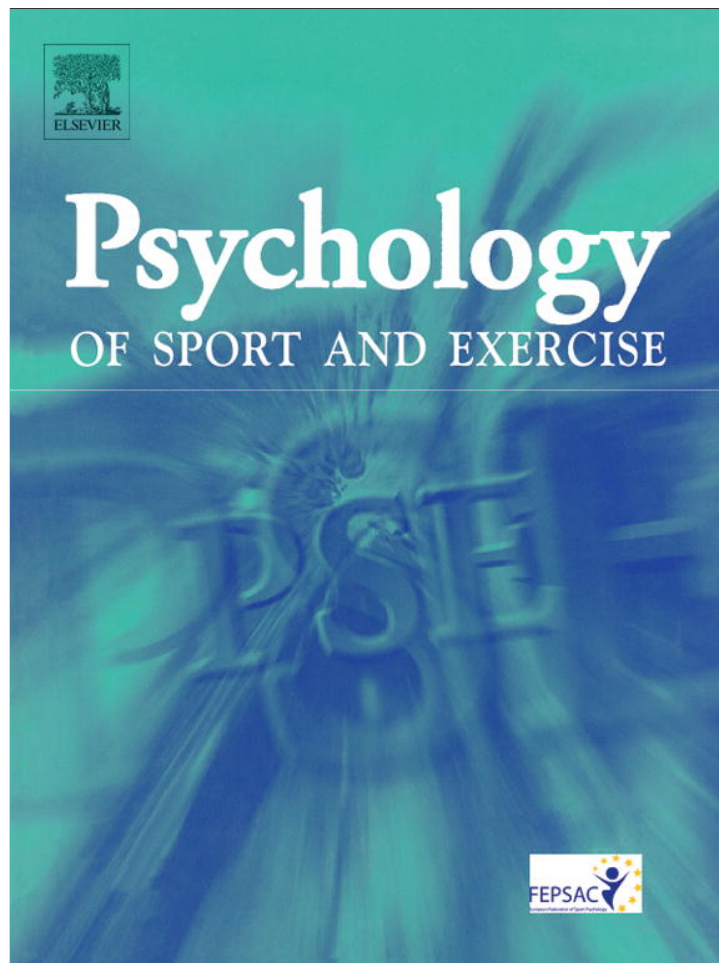


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

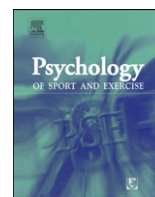
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Psychology of Sport and Exercise

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

Do the benefits from autonomy-supportive PE teacher training programs endure?: A one-year follow-up investigation

Sung Hyeon Cheon^a, Johnmarshall Reeve^{b,*}^aDepartment of Physical Education, Korea University, Seoul, Republic of Korea^bDepartment of Education, Korea University, Seoul, Republic of Korea

ARTICLE INFO

Article history:

Received 3 August 2012

Received in revised form

1 February 2013

Accepted 1 February 2013

Available online 27 February 2013

Keywords:

Autonomy support

Motivation

Physical education

Teacher training

ABSTRACT

Objective: An earlier study (Cheon, Reeve, & Moon, 2012) showed wide-ranging benefits from a training program designed to help teachers be more autonomy-supportive toward students during PE instruction. The present study collected a follow-up data set to determine whether those earlier-observed benefits endured one year later.

Design: We used an experimentally-based 3-wave longitudinal design. The experimental group consisted of 8 PE teachers from the original teacher training study and their 470 middle- and high-school students; the control group consisted of 9 matched PE teachers and their 483 students. Dependent measures included 3 manipulation checks, 3 measures of student motivation, and 6 course-specific outcomes.

Method: Trained raters scored teachers' instructional behaviors at mid-semester, while students reported perceptions of their teachers' motivating style and their own course-related motivation and outcomes at the beginning, middle, and end of the semester. We tested our hypotheses using hierarchical linear modeling to account for the hierarchical structure of data in which repeated measures were nested within students who were nested within teachers.

Results: Compared to teachers in the control group, teachers in the experimental group were scored by raters and perceived by students as more autonomy supportive and less controlling. Their students consistently reported greater motivation and more positive outcomes than did the students of teachers in the control group. All 8 teachers in the experimental group reported being significantly more autonomy supportive than a year earlier.

Conclusion: Teacher- and student-related benefits from the earlier autonomy-supportive training program endured.

© 2013 Elsevier Ltd. All rights reserved.

Physical education (PE) teachers prepare learning objectives and activities for their students and they deliver that instruction through the interpersonal context of a motivating style. Motivating style involves the tone of the teacher's sentiment and behavior while trying to motivate and engage students during instruction; it can be characterized within a bipolar continuum that extends from a highly controlling through a neutral style to one that is highly autonomy supportive (Deci, Schwartz, Sheinman, & Ryan, 1981; Reeve, 2009). When autonomy supportive, PE teachers motivate and engage students by adopting their perspective, by inviting, welcoming, and incorporating students' thoughts, feelings, and behaviors into the flow of instruction, and by supporting students' capacity for autonomous self-regulation, such as when they say, "Here is an opportunity to learn

something about yourself—what do you think of it?"; when controlling, PE teachers motivate and engage students by pressuring them into thinking, feeling, and behaving in a teacher-prescribed way, such as when they say, "Hurry. Do it this way—just like I showed you. Now!" (Assor, Kaplan, & Roth, 2002; Assor, Kaplan, Kanat-Maymon, & Roth, 2005; Reeve, 2009; Reeve, Jang, Carrell, Jeon, & Barch, 2004). Motivating style is important because it predicts course-related outcomes, as students with autonomy-supportive PE teachers, compared to those with controlling PE teachers, show greater autonomous motivation, classroom engagement, physical activity, performance, and achievement (Chatzisarantis & Hagger, 2009; Cheon et al., 2012; Moustaka, Vlachopoulos, Kabitsis, & Theodorakis, 2012; Vansteenkiste, Simons, Soenens, & Lens, 2004).

Empirical research rather strongly supports two conclusions: (1) PE teachers can learn how to become significantly more autonomy supportive toward students during classroom instruction; and (2) the students of these trained PE teachers benefit in a variety of important ways (Chatzisarantis & Hagger, 2009; Cheon et al., 2012; Tessier, Sarrazin, & Ntoumanis, 2008, 2010). Together, these two

* Corresponding author. Brain and Motivation Research Institute (bMRI), 633 Uncho-Useon Hall, Department of Education, Anam-Dong, Seongbuk-Gu, Korea University, Seoul 136-701, Republic of Korea.

E-mail address: reeve@korea.ac.kr (J. Reeve).

empirical conclusions show that PE teachers can, through their participation in teacher training programs, (1) transform their classroom motivating styles rather markedly away from traditionally controlling or neutral styles toward a highly autonomy-supportive style and, by doing so, (2) provide their students with a meaningfully-upgraded course experience that yields improved class functioning (e.g., greater engagement) and course outcomes (e.g., greater intentions for future physical activity).

To substantiate these claims, Table 1 lists the seven published studies that have carried out autonomy-supportive training programs (interventions) in the field of exercise promotion and physical education. Each study produced a significant treatment effect, and the table highlights the teacher- and student-based benefits (i.e., dependent measures) observed in each study. Though the studies varied in their samples, duration of training, research design, and dependent measures, they collectively support the conclusion that the interventions have been successful. Yet, none of the studies listed in Table 1 followed up these teachers after their participation in the autonomy-supportive intervention to assess the potential long-term benefits of the teacher-training program.

The question driving the present study was whether or not these positive training-induced benefits would endure if reassessed one year later. This represents a crucial question in assessing the efficacy of these training programs because exercise instructors and classroom PE teachers generally receive a wealth of support during these interventions that is then discontinued. During the teacher training experience, participating teachers observe and interact with expert models in how to be autonomy supportive, set classroom goals to improve their instructional strategies, receive guidance and feedback on their progress to enact autonomy-supportive instructional behaviors, complete reflection-facilitating teaching diaries, and participate in group discussions to

voice their concerns, identify potential obstacles, and share ideas and possible instructional strategies with their peers. Given such support, participants consistently have been able to revise their typically neutral or controlling motivating styles to become more autonomy supportive toward students. It remains an open question, however, as to whether the positive benefits of the training endure once the intensive support system is removed and teachers are effectively left on their own to instruct a new group of students.

To address our research question, we planned a one-year follow-up investigation. In the earlier study (Cheon et al., 2012), PE teachers collectively showed a wide-range of benefits after receiving a state-of-the-art intervention program. The teacher-training program was delivered over the course of an 18-week semester in three parts. In Part 1, PE teachers received a workshop experience on the nature of student motivation and teachers' motivating styles (i.e., what they are, where they come from, what outcomes they predict), classroom examples of autonomy-supportive instructional behavior, empirical evidence on the benefits of teacher-provided autonomy support, and a group discussion about the feasibility of, potential obstacles to, and recommendations on how to support students' autonomy during PE instruction. Part 2 took place 6 weeks later and after teachers had a first-hand opportunity to practice classroom-based autonomy support on a daily basis with their own students. It revolved around a group discussion that shared, critiqued, and refined PE-specific autonomy-supportive instructional strategies. Part 3 took place 6 weeks later and it too consisted of a group discussion centered around sharing and exchanging ideas on how to be autonomy-supportive during PE instruction. Teachers also completed a weekly journal activity.

The original Cheon et al. (2012) teacher-training program took place from February to early-July 2010 (semester 1 in the Korean school system). The plan of the present study was to revisit these

Table 1

Summary of all training intervention studies carried out with physical education teachers or exercise instructors on how to be more autonomy supportive toward students.

Reference citation	Sample	Duration of training	Research design	Benefits to teachers	Benefits to students
Chatzisarantis and Hagger (2009)	10 High-school PE teachers and their 215 students	5 Weeks	Quasi-experimental	Greater AS	Greater autonomous motivation Greater intention for PA Greater leisure time PA behavior
Cheon and Moon (2010)	1 Exercise instructor and his 60 university students	6 Weeks	Quasi-experimental	Greater AS	Greater autonomy, competence, and relatedness Greater autonomous motivation Greater future intentions for PA
Cheon et al. (2012)	19 Middle- and high-school PE teachers and their 1158 students	13 Weeks	Experimental	Greater AS Lesser CT Greater ASIBs	Greater autonomy, competence, and relatedness Greater autonomous motivation Lesser amotivation Greater classroom engagement Greater skill development Greater future intentions for PA Greater academic achievement
Edmunds, Ntoumanis, and Duda (2008)	1 Exercise instructor and her 56 university students taking an exercise class	9 Weeks	Quasi-experimental	Greater AS	Greater competence, relatedness Greater positive affect Greater exercise class attendance
Moustake, Vlachopoulos, Kabitsis, & Theodorakis (2012)	1 Exercise instructor and her 35 women taking a community-based exercise class	8 Weeks	Quasi-experimental	Greater AS	Greater autonomy, competence Greater autonomous motivation Lesser amotivation Greater vitality Greater exercise class attendance Greater exercise participation
Tessier et al. (2008)	5 Middle- and high-school PE teachers and their 96 students	8 Weeks	Experimental	Greater AS Greater verbal praise	n/a
Tessier et al. (2010)	3 Secondary PE teachers and their 185 students	4 Weeks	Pretest vs. posttest	Greater AS	Greater relatedness Lesser external regulation Lesser amotivation Greater collective engagement

Note. PE = physical education; AS = autonomy-supportive teaching; PA = physical activity; CT = controlling teaching; ASIBs = autonomy supportive instructional behaviors.

same previously-trained teachers one year later to see if (1) they were or were not still autonomy supportive toward students and (2) their (new) students perceived them as autonomy supportive and reported experiencing the same set of benefits as did the 2010 classes. The plan was to re-assess all the same dependent measures from the year-earlier study but in a new data set. The original Cheon et al. study assessed 12 dependent measures (3 manipulation checks, 3 student motivations, and 6 student outcomes), and we asked a new group of raters and a new group of students to report their perceptions of these teachers' motivating style and their own course-related motivation and outcomes. In addition, we added a new 2-item post-study follow-up questionnaire for the teachers in the experimental group that asked if their motivating style had changed compared to last year and, if so, why.

The global hypothesis across all 12 dependent measures was that the benefits observed in the original study would be replicated one year later—that is, we hypothesized that the earlier-observed teacher and student benefits would endure, even though teachers in the experimental group received no additional (follow-up) training in how to be autonomy supportive. We further hypothesized that teachers would report that they were either equally or more autonomy-supportive now than they were a year later. That said, we recognize the possibility that these teachers might revert back to their pre-existing motivating styles in the absence of formal support and in the presence of daily pressures to be controlling, such as time pressures, teacher accountability for externally-prescribed behavior and outcomes, and the press for immediate solutions to problems such as student misconduct and non-participation (Pelletier, Seguin-Lévesque, & Legault, 2002; Taylor, Ntoumanis, & Smith, 2009; Taylor, Ntoumanis, & Standage, 2008). Still, despite the absence of formal support and the presence of daily pressure, the previous training experience did allow participating teachers to become aware of—often for the first time—the benefits of autonomy support and the costs of teacher control. The autonomy-supportive motivating style they developed and refined during the year-earlier training led to easily observable student benefits—an assertion we make because the effect sizes on all student outcomes reported in Cheon et al. (2012) were so large as to be obvious classroom occurrences. Witnessing these benefits, these trained teachers might then decide to sustain an autonomy-supportive motivating style on the belief that its practice would reproduce these benefits for their new group of students. So, our predictions were that these previously-trained teachers would still be highly autonomy-supportive one-year later, that their new groups of students would show a very positive profile of PE-specific motivation and outcomes (compared to a control group of students), and that they would rate themselves either as equally autonomy supportive or as more autonomy supportive as a year earlier.

Method

Participants and procedure

Teacher participants

The original Cheon et al. (2012) study involved 19 PE teachers who completed all aspects of the study—10 in the experimental group and 9 in a delayed-treatment control group. For the present study, all 10 PE teachers in the original experimental condition received an invitation to participate in the present follow-up study. Eight of these teachers agree to participate in the follow-up study, while two teachers declined the invitation. As to the two teachers who declined the invitation, both reported the same reason for doing so—namely, because each teacher had just changed to a new school. In the Korean public school system, all teachers rotate to a

new school assignment every four or five years, and such a move means the affected teacher will be asked to take on new and highly time-consuming administrative duties.¹ The two dropout teachers did not differ from the eight persisting teachers on any teacher demographic characteristic or student dependent measure collected in the original 2010 study, all $t_s < 1$. The eight persisting teachers (3 women, 5 men) taught in eight different schools (6 middle schools, 2 high schools) within the Seoul, South Korea metropolitan area. Each teacher taught between five and eight classes, with class sizes that ranged from 35 to 40 students. The teachers had an average of 6.4 years of teaching experience (range = 3–9 years) and were, on average, 31.3 years of age (range = 27–34). The day-to-day content of the PE course curriculum they taught was prescribed and standardized by the Korean National and Educational Curriculum (KNEC) and revolved around week-long exposure to a series of different sport activities (e.g., badminton, rope jumping, basketball, track and field, table tennis).

To create the control group, we could not invite the teachers from the original Cheon et al. study to participate in the follow-up investigation, because all these teachers received the training intervention as a delayed-treatment during semester 2 in 2010. Instead, drawing on a large group of potential teacher-participants, our strategy was to create a matched-group of PE teachers by pairing each teacher in the experimental group with a matched control PE teacher in terms of gender, grade level taught, class size, age, teaching experience, and geographical location of the school. We invited 10 matched teachers, and 9 agreed to participate in the study, including 3 women and 6 men. These teachers had an average of 6.3 years of teaching experience (range = 3–11 years) and were, on average, 34.0 years of age (range = 29–40). They taught the same KNEC-prescribed course content as did teachers in the experimental group.

All 17 teacher participants (8 in the experimental group, 9 in the control group) completed all three waves of data collection, which included (1) having their students complete the study questionnaire at the beginning (week 1), middle (week 8), and end of the semester (week 17 or 18) and (2) allowing trained raters to score their instructional behavior in the middle of the semester (week 10). No teacher in the present study dropped out over the three-waves of data collection (i.e., the teacher retention rate was 100%). All 17 teachers were ethnic Korean, and all received the equivalent of \$50 in appreciation of their participation.

Student participants

The student participants who consented to complete the study questionnaire during the first week of classes (T1) were 1075 ethnic Korean students. During the second wave of data collection, 995 of the original 1075 student participants agreed to complete the questionnaire (retention rate = 92.6%). The 995 persisting students from T1 did not differ significantly from the 80 dropout students from T2 on any of the student-assessed T1 dependent measures, all $t_s < 1$, a result that suggests that student drop out occurred for random, rather than for systematic, reasons. During the third wave of data collection, 953 of the 995 students from the first two waves of data collection agreed to complete the questionnaire. The 953 persisting student participants from T3 did not differ significantly from the 42 T3 dropouts on any of the student-assessed T1 or T2 dependent measures, all $t_s < 1$. This final sample of 953 student

¹ With each move, teachers not only need to become familiar with new facilities, students, colleagues, and administrators, but they also transition from a teaching position with high seniority to one with low seniority. Most of the department-related administrative duties are done by the low seniority faculty. Knowing that their move would require a commitment to these time-consuming duties, our two dropout teachers told us that it was necessary for them to decline our invitation to participate in the follow-up study.

participants represented a retention rate of 88.7% (953/1075) and consisted of the following: 467 (49%) females and 486 (51%) males; 687 (72%) middle-school and 266 (28%) high-school students; and 470 (49%) students in the experimental and 483 (51%) in the control group.

Objective raters and their rating sheet

Before the study, four trained raters who were familiar both with PE instruction in Korean secondary schools and with the self-determination theory framework on teachers' motivating styles were trained in how to score teachers' classroom instruction in terms of autonomy-supportive and controlling instructional behaviors. For their training, they receive conceptual definitions of autonomy-supportive and controlling teaching, became familiar with a previously-validated rating sheet to operationally define autonomy-supportive and controlling teaching, received modeling and guidance in how to use the rating sheet (e.g., always start your rating at the midpoint score—a 4 on the 1–7 rating—and move away from that midpoint in response to what the teacher says and does during instruction), practiced using the rating sheet for two weeks by observing and scoring PE teachers first through videotaped instruction and then during live classroom instruction, and engaged in recurring in-rating and post-rating discussions with the authors to (a) explain, defend, and refine their ratings, (b) generate a single rating from 50 min of classroom observation for each of the five instructional behaviors, and (c) enhance their likelihood of producing high inter-rater reliabilities.

As illustrated graphically in Fig. 1, these raters visited the PE classes of teachers in both the experimental and control groups during week 10 of the 18-week semester. In doing so, they worked in pairs, came to the class unannounced 5 min before the start of the class, and did not know into which group (experimental or control) the observed teacher had been assigned. The two raters made independent ratings as they non-intrusively scored the classroom dynamics. The rating sheet was the same as that used in Cheon et al. (2012) (see Fig. 2, p. 372) that listed the following four autonomy-supportive instructional behaviors: nurtures inner motivational resources; relies on noncontrolling and informational language; provides explanatory rationales; and acknowledges and accepts negative affect. Each instructional behavior was scored using a bipolar format in which the controlling behavior (scored as 1) with illustrative descriptors appeared on the left side of the scoring sheet while the autonomy-supportive behavior (scored as 7) with illustrative descriptors appeared on the right side. Following the recommendation of Reeve (2009), we added a fifth autonomy-supportive instructional behavior to the rating sheet: Displays patience to allow for self-paced learning. For this item, "Displays Impatience" appeared on the left side of the rating sheet described by "Rushes student to produce the right answer, solution, or behavior" and "Communicates there is a right answer and the

student needs to reproduce it", while "Displays Patience" appeared on the right side described by "Allows student time and space for self-paced learning" and "Allows student to work in their own way and at their own pace". In the current study, the ratings from the two observers were highly positively correlated on each of the five instructional behaviors [range of five *r*'s (17) = 0.87–0.93]. Given these high reliabilities, we averaged the two raters' scores to produce the five objectively-scored ratings of teachers' autonomy-supportive versus controlling instructional behaviors.

Student self-report measures

Student participants completed the same four-page questionnaire used in the original Cheon et al. (2012) study to assess the same two manipulation checks (perceptions of autonomy-supportive teaching, perceptions of controlling teaching), three student motivations (autonomy, competence, and relatedness need satisfaction), and six course outcomes (autonomous motivation, amotivation, classroom engagement, perceived skill development, intentions toward future physical activity, and PE-specific course achievement). The psychometric properties and prior successful use of the questionnaires used to score these dependent measures within the exercise and PE literature are reported in detail in the Cheon et al. (2012) article. Though the measures were originally developed in English, the questionnaire as a whole had been previously back-translated into Korean and successfully used, as reported in Cheon et al. (2012). Each measure utilized the same 7-point Likert scale that ranged from 1 (*strongly disagree*) to 7 (*strongly agree*).

The name of each scale, its original reference citation, its number of items, the alpha coefficients obtained in the present study at T1, T2, and T3, and a sample item for each of the 17 measures embedded within the student questionnaire appear in Table 2. The measures for autonomy, competence, and relatedness need satisfaction, amotivation, perceived skill development, and future intentions were used singly as their own dependent measures. The scales for behavioral, emotional, cognitive, and agentic engagement were equally weighted and averaged into the single dependent measure of "classroom engagement". The measures for intrinsic motivation (IM), identified regulation (ID), introjected regulation (IJ), and extrinsic motivation (EX) were combined into the single dependent measure of "autonomous motivation" using the weighted formula recommended by Goudas, Biddle, and Fox (1994): $2(IM) + ID - IJ - 2(EX)$. Anticipated achievement was scored using the single item shown in Table 2.

Qualitative teacher reports

Two weeks after the semester ended, we asked the eight teacher participants in the experimental group to complete a 2-item questionnaire. The first item asked, "Compared to last year when

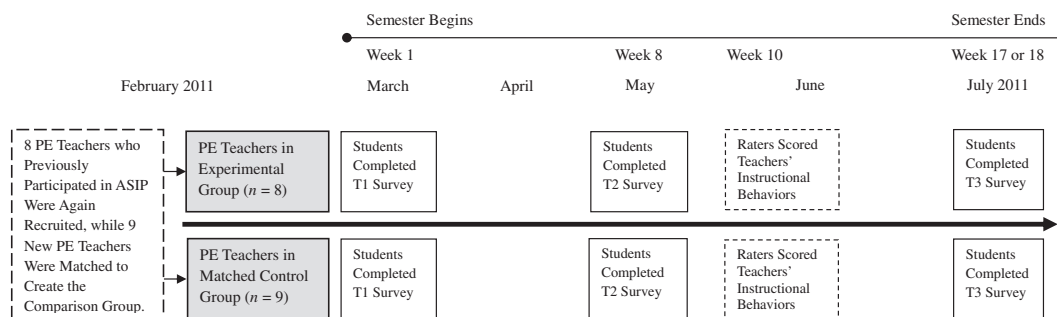


Fig. 1. Procedural timeline for the data collection.

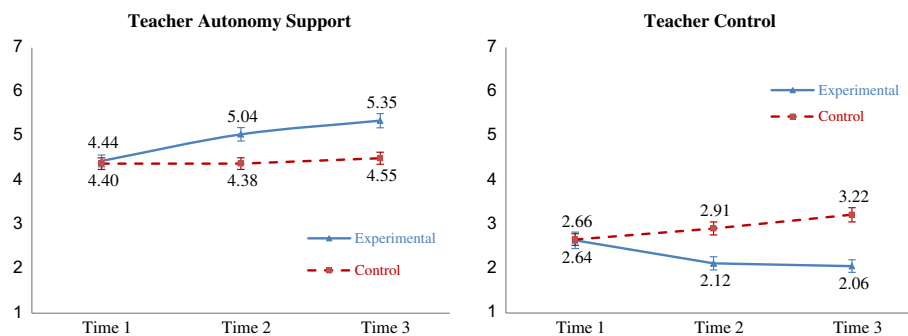


Fig. 2. Students' perceived teacher autonomy support (left panel) and perceived teacher control (right panel) broken down by experimental condition and time of assessment. *Note.* Numbers are mean scores, while the vertical bars represent the standard errors of those means. Solid lines and triangles represent the experimental group, while dashed lines and boxes represent the control group.

you completed the informational session on how to be autonomy supportive toward your students, would you say that you, this year, were more autonomy supportive, less autonomy supportive, or about the same in terms of autonomy support?" Teacher participants were asked to check one of the following three response options: more autonomy supportive than last year; about the same

as last year; less autonomy supportive than last year. The second item was an open-ended follow-up to the first question, "What reason or reasons explain why you checked the option you checked in question #1?" Instead of formally scoring responses to this second question, we sought insight into why teachers did or did not continue to support their students' autonomy during instruction.

Table 2 Name, original reference citation, number of items, alpha coefficients (T1, T2, and T3), and sample item for each measure on student questionnaire.

Dependent measure	Name of scale	Reference citation for original scale	Number of items	Alpha coefficient			Sample item
				T1	T2	T3	
Perceived autonomy support	Learning climate questionnaire	Williams and Deci (1996)	6	0.87	0.90	0.92	My PE teacher provides me with choices and options.
Perceived teacher control	Teacher control questionnaire	Jang, Reeve, Ryan, and Kim (2009)	4	0.81	0.86	0.88	My PE teacher puts a lot of pressure on me.
Autonomy need satisfaction		Standage, Duda, and Ntoumanis (2006)	5	0.85	0.89	0.89	In this PE class, I feel that I do PE activities because I want to.
Competence need satisfaction	Intrinsic motivation inventory	McAuley, Duncan, and Tammen (1989)	4	0.90	0.90	0.89	I think I am pretty good at PE.
Relatedness need satisfaction		Furrer and Skinner (2003)	4	0.77	0.82	0.83	When I am with my PE teacher, I feel accepted.
Behavioral engagement		Skinner, Kindermann, and Furrer (2009)	5	0.85	0.87	0.87	In PE class, I work as hard as I can.
Emotional engagement		Skinner et al. (2009)	5	0.89	0.90	0.80	When I am in this PE class, I feel good.
Cognitive engagement		Wolters (2004)	3	0.85	0.88	0.88	When learning a PE activity, I try to relate what I am learning to what I...
Agentic engagement	Agentic engagement scale	Reeve and Tseng (2011)	5	0.81	0.86	0.87	During this PE class, I express my preferences and opinions.
Intrinsic motivation	Perceived locus of causality scale	Goudas et al. (1994)	4	0.89	0.91	0.91	I take part in this PE class, because it is enjoyable.
Identified regulation	Perceived locus of causality scale	Goudas et al. (1994)	4	0.84	0.87	0.90	I take part in this PE class, because it is important.
Introjected regulation	Perceived locus of causality scale	Goudas et al. (1994)	4	0.60	0.65	0.65	I take part in this PE class, because I want the teacher to think I am a good student.
Extrinsic regulation	Perceived locus of causality scale	Goudas et al. (1994)	4	0.53	0.54	0.51	I take part in this PE class, because that is the rule.
Amotivation	Perceived locus of causality scale	Goudas et al. (1994)	4	0.91	0.90	0.91	I take part in this PE class, but I can't see what I am getting out of PE.
Perceived skill development	Perceived skill in PE	Cheon et al. (2012)	5	0.91	0.93	0.93	I have learned new and important skills during this PE class.
Intentions for future physical activity		Ntoumanis (2005)	3	0.83	0.87	0.87	In the future, I intend to make sports and physical activity a part of my life.
Anticipated achievement	Anticipated academic achievement	Jeon (2007)	1	—	—	—	In this PE class, I expect that my course grade will be ---- points (enter a number between 0 and 100).

Results

Preliminary analyses

Before testing the hypotheses, we conducted multilevel analyses using hierarchical linear modeling (HLM, Version 7.0; Raudenbush, Bryk, & Congdon, 2011) to determine whether meaningful between-teacher differences might have affected the student-reported dependent measures. The intra-class correlation coefficients (ICCs) associated with the 11 T1 student-assessed dependent measures calculated from unconditional models were as follows: perceived autonomy support, 14.8%; perceived controlling, 8.8%; autonomy need satisfaction, 8.1%; competence need satisfaction, 3.3%; relatedness need satisfaction, 7.2%; autonomous motivation, 10.9%; amotivation, 15.1%; classroom engagement, 9.7%; perceived skill development, 8.9%; future intentions, 7.6%; and anticipated achievement, 3.6%. Given these meaningful between-teacher effects, we used multilevel modeling to represent the nested nature of the data—namely, longitudinal data collected from students nested within teachers. By doing so, we sought to partial out the “between-teacher” effects within the students’ data (as represented by the ICCs that averaged 8.9% across the 11 dependent measures) such that the analyses tested the hypotheses in a way that students’ scores on each dependent measure were statistically independent of these “controlled for” teacher-level effects.

The present longitudinal design had a three-level hierarchical structure with repeated measures (Level 1) nested within students (Level 2) nested within teachers (Level 3). At level 1 (within student), the longitudinal data allowed us to study students’ increase or decrease on each dependent measure over three time points—the beginning, middle, and end of the semester. We centered the “time” independent variable on participants’ beginning-of-semester score so that T1 scores functioned as an initial status measurement on each dependent measure so that the T2 and T3 scores functioned as change scores from that initial status score. At level 2 (between students), we entered the student-level individual difference of gender as a time- or wave-invariant covariate to function as a statistical control to recognize the “gender gap” observed in the earlier Cheon et al. (2012) study in which males showed a more motivationally constructive profile in PE class than did females. Gender was group-mean centered in all analyses. At level 3 (between teachers), we entered experimental condition as an un-centered independent variable so that we could retain its raw metric form of experimental group = 1 and control group = -1. Finally, we entered the crucial condition \times time interaction as a cross-level predictor to test the extent to which the changes in the T2 and T3 scores on each dependent measure depended on experimental condition, that is, on whether the student’s teacher was in the experimental or control group.

Manipulation checks

We assessed the extent to which the two groups of teachers differed in their provision of autonomy-supportive instruction in two ways. First, trained raters scored teachers’ objective classroom instructional behaviors at mid-semester. Second, students reported on their teachers’ perceived autonomy-supportive and perceived controlling teaching at three times during the semester (T1, T2, and T3).

Raters’ objective scoring of motivating style

Raters scored teachers in the experimental group as enacting significantly more autonomy-supportive instructional behaviors than the teachers in the control group: nurtures inner motivational

resources, Ms, 6.13 vs. 4.42; $t(15) = 3.78, p < 0.01, d = 0.98$; uses informational language, Ms, 6.53 vs. 4.28; $t(15) = 5.68, p < 0.01, d = 1.47$; offers explanatory rationales, Ms, 6.47 vs. 3.92; $t(15) = 8.44, p < 0.01, d = 2.18$; acknowledges and accepts expressions of negative feelings, Ms, 6.28 vs. 4.06; $t(15) = 5.60, p < 0.01, d = 1.45$; and displays patience, Ms, 6.41 vs. 4.86; $t(15) = 4.05, p < 0.01, d = 1.05$. The average effect size across these instructional behaviors was $d = 1.43$.

Students’ perceptions of motivating style

To assess the effect of experimental condition on students’ perceptions of their teachers’ motivating style, we conducted HLM-based multi-level regression analyses in which time of assessment (T1, T2, and T3) was the within-student Level 1 repeated measure, experimental condition was the between-teacher Level 3 hypothesized predictor, and gender was the between-student Level 2 covariate (statistical control). The test of each hypothesis was for a significant condition \times time/wave interaction. Mean scores for autonomy-supportive and controlling teaching, adjusted for the gender covariate and between-teacher effects, appear in Fig. 2 broken down by experimental condition and time of assessment. In conducting pair-wise mean comparisons, we used the Bonferroni corrected t -test procedure (family-wise $\alpha = 0.05/6 = 0.008$) to provide the alpha level ($\alpha = 0.008$) used in each of the six post hoc mean comparisons for each dependent measure. As expected, students’ perceptions of autonomy-supportive and controlling teaching were consistently negatively correlated; T1, $r(953) = -0.38, p < 0.01$; T2, $r(953) = -0.40, p < .01$; and T3, $r(953) = -0.38, p < 0.01$.

For perceived autonomy-supportive teaching, the condition main effect was only marginally significant, $t(15) = 1.95, p < 0.07$, the time main effect was significant, $t(1886) = 11.25, p < 0.01$, and, most importantly, the condition \times time interaction was significant, $t(1886) = 6.82, p < 0.01$. As illustrated in the left panel of Fig. 2, perceived autonomy support increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.60, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.31, p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = -0.02, ns$) but then surprisingly increased significantly from T2 to T3 ($\Delta = +0.17, p < 0.008$). While the two conditions did not differ at the T1 baseline ($\Delta = 0.04, ns$), perceived autonomy support was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.66, p < 0.008$) and T3 ($\Delta = +0.80, p < 0.008$).

For perceived controlling teaching, the condition main effect was significant, $t(15) = 2.63, p < 0.05$, the time main effect was not significant, $t(1886) = 0.40, ns$, and the condition \times time interaction was significant, $t(1886) = 9.00, p < 0.01$. As illustrated in the right panel of Fig. 2, perceived controlling teaching decreased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = -0.52, p < 0.008$) but then only leveled off from T2 to T3 ($\Delta = -0.06, ns$), while it increased significantly for students of the teachers in the control group from T1 to T2 ($\Delta = +0.25, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.31, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = 0.02, ns$), perceived controlling was lesser for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = -0.79, p < 0.008$) and T3 ($\Delta = -1.16, p < 0.008$).

Student psychological need satisfaction

For autonomy need satisfaction, the condition main effect was not significant, $t(15) = 1.74, ns$, the time main effect was significant, $t(1886) = 12.26, p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 5.86, p < 0.01$. As illustrated in the left panel of

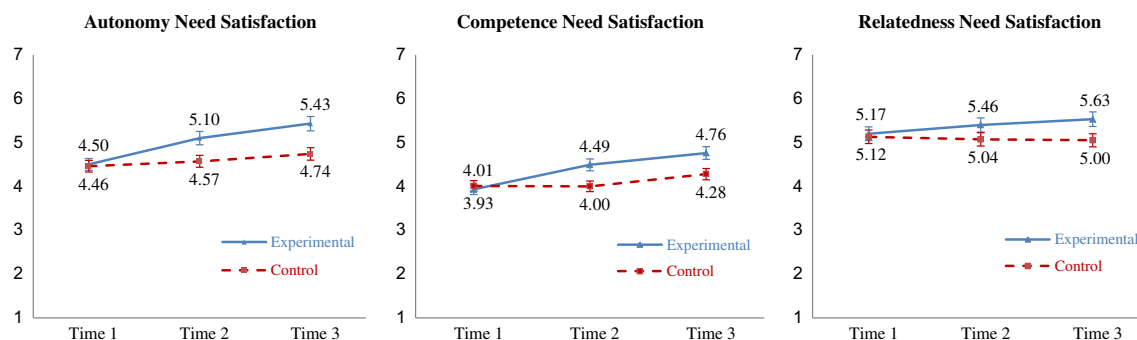


Fig. 3. Students' self-reports of autonomy (left panel), competence (center panel), and relatedness (right panel) psychological need satisfaction broken down by experimental condition and time of assessment. *Note.* Numbers are mean scores, while the vertical bars represent the standard errors of those means. Solid lines and triangles represent the experimental group, while dashed lines and boxes represent the control group.

Fig. 3, autonomy need satisfaction increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.60, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.33, p < 0.008$), while it also increased significantly for students of the teachers in the control group from T1 to T2 ($\Delta = +0.11, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.17, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = 0.04, ns$), autonomy need satisfaction was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.53, p < 0.008$) and T3 ($\Delta = +0.69, p < 0.008$).

For *competence need satisfaction*, the condition main effect was not significant, $t(15) = 0.89, ns$, the time main effect was significant, $t(1886) = 10.52, p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 3.79, p < 0.01$. As illustrated in the center panel of Fig. 3, competence need satisfaction increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.56, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.27, p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = -0.01, ns$) but then increased significantly from T2 to T3 ($\Delta = +0.28, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = -0.08, ns$), competence need satisfaction was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.49, p < 0.008$) and T3 ($\Delta = +0.48, p < 0.008$).

For *relatedness need satisfaction*, the condition main effect was only marginally significant, $t(15) = 1.81, p < 0.09$, the time main effect was significant, $t(1886) = 3.63, p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 5.92, p < 0.01$. As illustrated in the right panel of Fig. 3, relatedness need satisfaction increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.29, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.17, p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = -0.08, ns$) and was further unchanged from T2 to T3 ($\Delta = -0.04, ns$). While the two conditions did not differ at baseline ($\Delta = 0.05, ns$), relatedness need satisfaction was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.42, p < 0.008$) and T3 ($\Delta = +0.63, p < 0.008$).

Student outcomes

For *autonomous motivation*, the condition main effect was not significant, $t(15) = 1.46, ns$, the time main effect was not significant, $t(1886) = 1.02, ns$, and the condition \times time interaction was significant, $t(1886) = 7.74, p < 0.01$. As illustrated in the top-left panel of Fig. 4, autonomous motivation increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +1.83,$

$p < 0.008$) but then leveled off from T2 to T3 ($\Delta = -0.23, ns$), while it decreased significantly for students of the teachers in the control group from T1 to T2 ($\Delta = -1.38, p < 0.008$) and decreased significantly again from T2 to T3 ($\Delta = -0.55, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = -0.39, ns$), autonomous motivation was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +2.82, p < 0.008$) and T3 ($\Delta = +3.14, p < 0.008$).

For *amotivation*, the condition main effect was not significant, $t(15) = 1.25, ns$, the time main effect was significant, $t(1886) = 2.74, p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 6.66, p < 0.01$. As illustrated in the top-center panel of Fig. 4, amotivation was unchanged for students of the teachers in the experimental group from T1 to T2 ($\Delta = -0.05, ns$) but then decreased significantly from T2 to T3 ($\Delta = -0.19, p < 0.008$), while it increased significantly for students of the teachers in the control group from T1 to T2 ($\Delta = +0.42, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.17, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = 0.04, ns$), amotivation was lesser for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = -0.43, p < 0.008$) and T3 ($\Delta = -0.79, p < 0.008$).

For *classroom engagement*, the condition main effect was not significant, $t(15) = 1.68, ns$, the time main effect was significant, $t(1886) = 9.57, p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 4.46, p < 0.01$. As illustrated in the top-right panel of Fig. 4, classroom engagement increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.49, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.13, p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = +0.05, ns$) but then increased significantly from T2 to T3 ($\Delta = +0.10, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = 0.07, ns$), classroom engagement was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.51, p < 0.008$) and T3 ($\Delta = +0.54, p < 0.008$).

For *perceived skill development*, the condition main effect was not significant, $t(15) = 1.56, ns$, the time main effect was significant, $t(1886) = 10.18, p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 5.60, p < 0.01$. As illustrated in the bottom-left panel of Fig. 4, perceived skill development increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.60, p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.23, p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = -0.01, ns$) but then increased significantly from T2 to T3 ($\Delta = +0.14, p < 0.008$). While the two conditions did not differ at baseline ($\Delta = 0.02, ns$), perceived skill development was greater for students of the

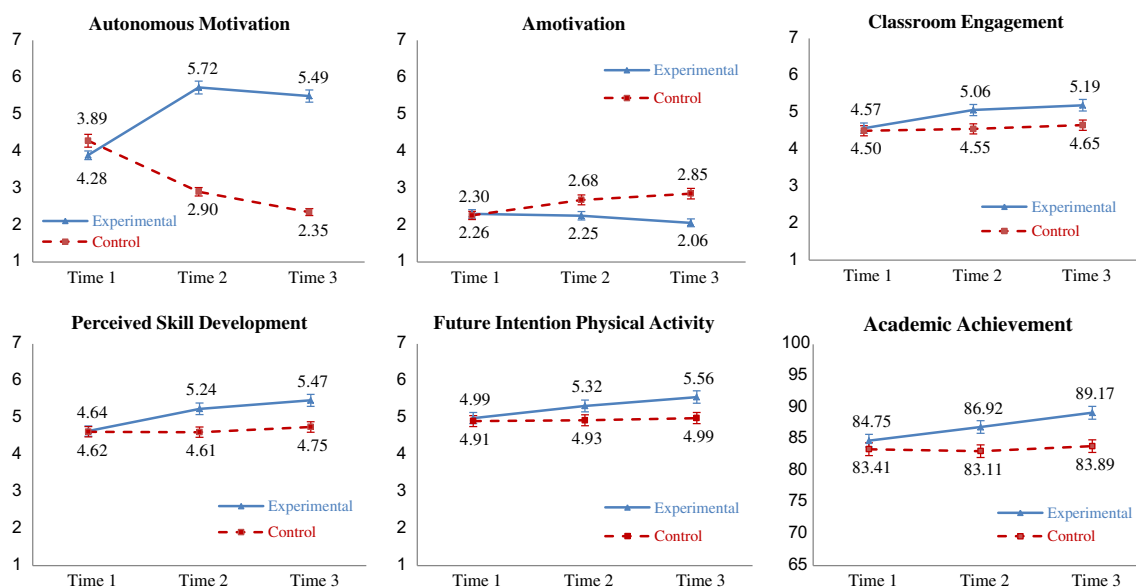


Fig. 4. Students' autonomous motivation (top-left panel), amotivation (top-center), classroom engagement (top-right), perceived skill development (bottom-left), future intentions (bottom-center), and anticipated achievement (bottom-right) broken down by experimental condition and time of assessment. *Note.* Numbers are mean scores, while the vertical bars represent standard errors of those means. Solid lines and triangles represent the experimental group, while dashed lines and boxes represent the control group.

teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.63$, $p < 0.008$) and T3 ($\Delta = +0.72$, $p < 0.008$).

For *future intentions*, the condition main effect was not significant, $t(15) = 1.28$, *ns*, the time main effect was significant, $t(1886) = 5.83$, $p < 0.01$, and the condition \times time interaction was significant, $t(1886) = 3.89$, $p < 0.01$. As illustrated in the bottom-center panel of Fig. 4, future intentions increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +0.33$, $p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +0.24$, $p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = +0.02$, *ns*) and was further unchanged from T2 to T3 ($\Delta = +0.06$, *ns*). While the two conditions did not differ at baseline ($\Delta = 0.08$, *ns*), future intentions were greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +0.39$, $p < 0.008$) and T3 ($\Delta = +0.57$, $p < 0.008$).

For *anticipated achievement*, the condition main effect was only marginally significant, $t(15) = 1.77$, $p < 0.10$, the time main effect was significant, $t(1886) = 3.36$, $p < 0.01$, and the condition \times time interaction was only marginally significant, $t(1886) = 1.86$, $p < 0.07$. As illustrated in the bottom-right panel of Fig. 4, anticipated achievement increased significantly for students of the teachers in the experimental group from T1 to T2 ($\Delta = +2.17$, $p < 0.008$) and increased significantly again from T2 to T3 ($\Delta = +2.25$, $p < 0.008$), while it was unchanged for students of the teachers in the control group from T1 to T2 ($\Delta = -0.30$, *ns*) and was further unchanged from T2 to T3 ($\Delta = +0.78$, *ns*). While the two conditions did not differ at baseline ($\Delta = 1.34$, *ns*), anticipated achievement was greater for students of the teachers in the experimental group than it was for students of teachers in the control group at both T2 ($\Delta = +3.81$, $p < 0.008$) and T3 ($\Delta = +5.28$, $p < 0.008$).

Qualitative teacher reports

When asked to compare their 2011 autonomy-supportive teaching to their 2010 autonomy-supportive teaching, all eight teachers answered the 3-option multiple choice question by reporting that they were "more autonomy supportive than last year", $\chi^2(2) = 15.98$,

$p < 0.001$. When asked why this was so, a male teacher cited both his greater familiarity with autonomy-supportive teaching and his students' positive response to it:

"At first it was strange but now it is familiar to me. It is now my way to meet and teach students in daily school life. Students had more questions and approached me in a friendly way when I welcomed them in an autonomy-supportive way, such as welcoming their feelings and thoughts, taking students' perspectives especially some who were unskilled. And I tried to be more autonomy supportive whenever my students truly wanted to do something, and I tried to respond to their actions by providing rationales and showing my patience."

A female teacher cited both observed student benefits and her own greater teaching confidence:

"The quality of physical education was enhanced when I supported students' autonomy. I felt happy and supported their autonomy when I found that students actually recognized what they truly valued and enjoyed. I was more confident in how to manage my students. Now, I always think before my class how to support students' autonomy."

The other six teachers in the experimental group all offered similar, though more lengthy, essays to explain their greater autonomy support. Consistently, each teacher emphasized the student benefits he or she observed and an improved quality in teacher–student relationships.

Discussion

We undertook the present study to determine whether or not the benefits observed in the original Cheon et al. study would endure when these same teacher- and student-related dependent measures were assessed one year later. Findings from the three teacher-related dependent measures and eight of the nine student-related dependent measures confirmed that the training-induced benefits did endure over time. Apparently, teachers used their earlier training experience to re-conceptualize their understandings of student motivation and of teachers' motivating

styles, to learn new autonomy-supportive ways of providing PE instruction, to develop and refine their skill in enacting autonomy-supportive instructional behaviors, and more generally to undertake a professional development opportunity to fundamentally change their approach toward motivating and engaging students during instruction.

These data suggest that the immediate training benefits observed in the seven autonomy-supportive teacher training programs summarized in Table 1 likely extend to long-term benefits as well. As expressed in the open-ended teacher quotations, the earlier autonomy-supportive teacher training program produced these enduring benefits partly because teachers observed positive changes in their students' motivation, engagement, and general responsiveness to classroom learning activities. Such student responsiveness seemed to feed back to confirm that teachers' newly-acquired instructional behaviors were indeed both effective and beneficial—and hence worthy of being integrated into their classroom motivating style in an enduring way. This suggests the interesting possibility that the autonomy-supportive teacher training program may provide teachers with an initial willingness to implement autonomy-supportive behaviors during instruction (though with a somewhat provisional “let's see if this works”) while students' positive responsiveness then confirms or validates that initial willingness to give it a try.

One potential problem within the design of the present study was the lack of random assignment to conditions. Random assignment is crucial if investigators want to create two groups of participants who are equal on the dependent measures at T1. Anticipating this potential problem, we intentionally chose to use a repeated-measures (i.e., longitudinal) analytic strategy to test each hypothesis. Doing so allowed us to account statistically for any T1 group-based differences that might have existed. That is, even if the experimental versus control groups of participants differed on a T1 dependent measure (because of the lack of random assignment), these T1 differences would be carried forward into the T2 and T3 assessments, because within-subject (repeated measures) analyses treat T1 differences as statistically controllable individual differences that test only for change in the dependent measure's initial status over time. The crucial hypothesis test is therefore to examine the condition \times time/wave interaction effect. For 10 of the 11 dependent measures (all but anticipated achievement, which was $p < 0.07$), this key interaction effect was significant. This pattern of results means that, after controlling for any T1 group differences (and also after controlling for any gender or between-teacher differences), the students of teachers in the experimental group consistently showed improved T2 and T3 dependent measures, compared to their counterparts in the control group.

One possible limitation to the present study—and to all the studies listed in Table 1—is that the dependent measures assessed only students' positive classroom functioning and course outcomes, though the inclusion of amotivation was the one exception. Of course, students taking PE courses sometimes display poor classroom functioning and negative course outcomes (e.g., psychological need thwarting, classroom disengagement, problematic relationships, and episodes of aggression). Recent research in the exercise promotion and PE classroom literature has shown the importance of students' frustrated motivational states and negative course outcomes (Aelterman et al., 2012; Bartholomew, Ntoumanis, Ryan, & Thøgersen-Ntoumani, 2011; Tessier et al., 2008). For instance, highly controlling PE instruction sometimes produces a large effect size on dependent measures such as autonomy, competence, and relatedness need frustration but only a small effect size on autonomy, competence, and relatedness need satisfaction. The implications are that (1) autonomy-supportive instructional behaviors might promote constructive motivation and

positive outcomes while controlling instructional behaviors might frustrate students' motivation and course outcomes and (2) constructive motivation and positive course outcomes versus frustrated motivation and negative course outcomes are not interchangeable opposites (e.g., need satisfaction is qualitatively different from need frustration). Recognizing this, a future state-of-the-art teacher training program might need to place just as much emphasis on recommending, modeling, and scaffolding against controlling teaching as it does on recommending, modeling, and scaffolding for autonomy-supportive teaching. One way to do this would be to suggest autonomy-supportive instructional behaviors replace controlling ones (e.g., offer challenge or curiosity-inducing questions to spark students' initial engagement, rather than rely on directives, commands, or extrinsic incentives). But this is sometimes a difficult transition for teachers to make, at least in a short period of time, because it represents a change in one's motivating style (from controlling to autonomy supportive) rather than an expansion to one's existing style (become more autonomy supportive). So, a second way to do this would be to suggest teachers provide highly-structured instruction (e.g., introduce a clear timeline to script students' forthcoming classroom activity). A follow-up step would then focus on learning how to introduce various elements of classroom structure (e.g., rules, expectations, feedback) in an autonomy-supportive way (for more on this recommendation, see Jang, Reeve, & Deci, 2010; Reeve, 2009).

A second possible limitation of the study was the use of bi-polar rating scales for the raters. We had our raters use bi-polar scales to score teachers' classroom instructional behaviors that treated autonomy support and controlling as opposite ends of a single continuum, following theoretical (Deci et al., 1981) and empirical (Reeve et al., 2004) precedent. We nevertheless recognize that other investigators prefer to assess autonomy-supportive and controlling instructional behaviors as separate categories; that is, teachers are scored once on how autonomy supportive they are but also, separately, on how controlling they are; further the correlations between these two ratings are sometimes low, which suggests that autonomy support and teacher control are somewhat independent aspects of motivating style, rather than opposites (Aelterman et al., 2012; Bartholomew et al., 2011; Tessier et al., 2008). The key difference between these two rating strategies is that, while our methodology rates teachers on categories of autonomy-supportive versus controlling instructional behavior (e.g., “uses informational versus pressuring language”) averaged over a 50 min period, the alternative methodology is to rate specific acts of instruction (e.g., the number of times the teacher says “you should”, the number of times the teacher “shouts or yells”) for briefer (e.g., 5 min) episodes. We had raters score teachers' instructional behaviors generally over the 50-min class period, rather than discretely for shorter intervals (e.g., a sequential series of 5- or 10-min rating periods), because our analyses of previous data sets (e.g., Reeve et al., 2004) showed that the two scoring strategies (one general 50-min rating vs. five aggregated 10-min ratings) produced functionally equivalent (interchangeable) data. That said, both the molar and molecular approaches have merit and can be appropriate for different types of research questions. A molecular approach (5- or 10-min rating periods) could look for within-class period trends in teachers' motivating styles, and it could also employ transactional models (Sameroff, 2009) to investigate how in-class changes in students' classroom motivation, engagement, cooperation, or performance might lead to in-class changes in teachers' instructional behaviors. In addition, a strategy to score autonomy-supportive and controlling instructional behaviors separately would allow for the possibility of constructing different teacher profiles (high on both types in instructional behaviors, low on both, high on one but low on the other).

A final limitation of the study was that, surprisingly, students of the teachers in the control group showed T2–T3 gains in 5 of the 11 student-scored dependent measures, including perceived autonomy support, autonomy and competence need satisfaction, classroom engagement, and perceived skill development. This was a surprise because the previous Cheon et al. (2012) study had not shown these same T2–T3 increases. We were unable to explain these T2–T3 changes for students of the teachers in the control group, as they might have occurred because all students—not just those with autonomy-supportive teachers—profited from the instruction they received to gain skill and a sense of competence in PE activities. Alternatively, the activities late in the semester (e.g., soccer) might have been more engagement-fostering than were those early in the semester (e.g., gate ball). That is, there might be a “type of activity” effect in these data that we did not measure. Another possibility might be that a teacher’s motivating style is more important (more predictive) earlier in the semester than it is later in the semester, a finding reported by Jang, Kim, and Reeve (2012). Late in the semester, changes in students’ own classroom engagement can function as an equally good predictor of students’ motivation and outcomes as does the teacher’s motivating style (Reeve, Lee, & Kim, submitted for publication). Irrespective of the reason underlying these T2–T3 increases, the key point evidenced by the significant condition \times time interaction effects was that the extent of the T2–T3 increase in these dependent measures was greater for students of teachers in the experimental group than it was for students of teachers in the control group.

Conclusion

The earlier-observed benefits teachers experienced from their participation in a state-of-the-art autonomy-supportive teacher training program endured one year later. This positive training effect likely endured because teachers used the earlier professional development experience as an opportunity to gain the knowledge, skill, and experience they needed to make a long-lasting and highly constructive change in how they went about the everyday classroom challenge of motivating and engaging their students during PE instruction.

Acknowledgments

This research was supported by the WCU (World Class University) Program funded by the Korean Ministry of Education, Science and Technology, consigned to the Korea Science and Engineering Foundation (Grant no. R32-2008-000-20023-0) and by the National Research Foundation of Korea funded by the Korean Government (Grant no. NRF-2012S1A5B5A07037749).

References

- Aelterman, N., Vansteenkiste, M., Van Keer, H., Van den Berghe, L., Meyer, J. D., & Haerens, L. (2012). Students’ objectively measured physical activity levels and engagement as a function of between-class and between-student differences in motivation toward physical education. *Journal of Sport & Exercise Psychology*, 34, 457–480.
- Assor, A., Kaplan, H., Kanat-Maymon, Y., & Roth, G. (2005). Directly controlling teacher behaviors as predictors of poor motivation and engagement in girls and boys: the role of anger and anxiety. *Learning and Instruction*, 15, 397–413.
- Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: autonomy-enhancing and suppressing teaching behaviors predicting students’ engagement in schoolwork. *British Journal of Educational Psychology*, 27, 261–278.
- Bartholomew, K. J., Ntoumanis, N., Ryan, R., & Thøgersen-Ntoumani, C. (2011). Psychological need thwarting in the sport context: development and initial validation of a psychometric scale. *Journal of Sport & Exercise Psychology*, 33, 75–102.
- Chatzisarantis, N. L., & Hagger, M. S. (2009). Effects of an intervention based on self-determination theory on self-reported leisure-time physical activity participation. *Psychology and Health*, 24, 29–48. <http://dx.doi.org/10.1080/08870440701809533>.
- Cheon, S. H., & Moon, I. S. (2010). Implementing an autonomy-supportive fitness program to facilitate students’ autonomy and engagement. *Korean Journal of Sport Psychology*, 21, 175–195.
- Cheon, S. H., Reeve, J., & Moon, I. S. (2012). Experimentally based, longitudinally designed, teacher-focused intervention to help physical education teachers be more autonomy supportive toward their students. *Journal of Sport and Exercise Psychology*, 34, 365–396.
- Deci, E. L., Schwartz, A., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adult’s orientations toward control versus autonomy in children: reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, 73, 642–650. <http://dx.doi.org/10.1037/0022-0663.73.5.642>.
- Edmunds, N. (2005). A prospective study of participation in optional school physical education using a self-determination theory framework. *Journal of Educational Psychology*, 97, 444–453. <http://dx.doi.org/10.1037/0022-0663.97.3.444>.
- Edmunds, J., Ntoumanis, N., & Duda, J. L. (2008). Testing a self-determination theory-based teaching style intervention in the exercise domain. *European Journal of Social Psychology*, 38, 375–388. <http://dx.doi.org/10.1002/ejsp.463>.
- Furrer, C., & Skinner, E. A. (2003). Sense of relatedness as a factor in children’s academic engagement and performance. *Journal of Educational Psychology*, 95, 148–162. <http://dx.doi.org/10.1037/0022-0663.95.1.148>.
- Goudas, M., Biddle, S., & Fox, K. (1994). Perceived locus of causality, goal orientations, and perceived competence in school physical education classes. *The British Journal of Educational Psychology*, 64, 453–463.
- Jang, H., Kim, E. J., & Reeve, J. (2012). Longitudinal test of self-determination theory’s motivation mediation model in a naturally occurring classroom context. *Journal of Educational Psychology*, 104, 1175–1188. <http://dx.doi.org/10.1037/a0028089>.
- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: it is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, 102, 588–600. <http://dx.doi.org/10.1037/a0019682>.
- Jang, H., Reeve, J., Ryan, R. M., & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically-oriented South Korean adolescents? *Journal of Educational Psychology*, 101, 644–661. <http://dx.doi.org/10.1037/a0014241>.
- Jeon, S. (2007). The effects of parents’ and teachers’ motivating styles on adolescents’ school outcomes and psychological well-being: a test of self-determination theory in a Korean context. Unpublished dissertation, University of Iowa, Iowa City, IA.
- McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: a confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, 60, 48–58.
- Moustaka, F. C., Vlachopoulos, S. P., Kabitsis, C., & Theodorakis, Y. (2012). Effects of an autonomy-supportive exercise instructing style on exercise motivation, psychological well-being, and exercise attendance in middle-age women. *Journal of Physical Activity and Health*, 9, 139–150.
- Pelletier, L. G., Seguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers’ motivation and teaching behaviors. *Journal of Educational Psychology*, 94, 186–196. <http://dx.doi.org/10.1111/j.2044-8279.1994.tb01116.x>.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, Jr. R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling [Computer software]*. Lincolnwood, IL: Scientific Software International, Inc.
- Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, 44, 159–178. <http://dx.doi.org/10.1080/00461520903028990>.
- Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students’ engagement by increasing teachers’ autonomy support. *Motivation and Emotion*, 28, 147–169.
- Reeve, J., Lee, W., & Kim, H. Students as architects of their own motivation: changes in engagement produce longitudinal changes in motivation, submitted for publication.
- Reeve, J., & Tseng, C. (2011). Agency as a fourth aspect of students’ engagement during learning activities. *Contemporary Educational Psychology*, 36, 257–267. <http://dx.doi.org/10.1016/j.cedpsych.2011.05.002>.
- Sameroff, A. (2009). *The transactional model of development: How children and contexts shape each other*. Washington, DC: American Psychological Association.
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: conceptualization and assessment of children’s behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, 69, 493–525. <http://dx.doi.org/10.1177/0013164408323233>.
- Standage, M., Duda, J. L., & Ntoumanis, N. (2006). Students’ motivational processes and their relationship to teacher ratings in school physical education: a self-determination theory approach. *Research Quarterly for Exercise and Sport*, 77, 100–110. <http://dx.doi.org/10.5641/027013606X13080769704046>.
- Taylor, I. M., Ntoumanis, N., & Smith, N. (2009). The social context as a determinant of teacher motivational strategies in physical education. *Psychology of Sport & Exercise*, 10, 235–243.
- Taylor, I. M., Ntoumanis, N., & Standage, M. (2008). A self-determination theory approach to understanding the antecedents of teachers’ motivational strategies in physical education. *Journal of Sport and Exercise Psychology*, 30, 75–94.
- Tessier, D., Sarrazin, P., & Ntoumanis, N. (2008). The effects of an experimental programme to support students’ autonomy on the overt behaviours of physical education teachers. *European Journal of Psychology of Education*, 23, 239–253. <http://dx.doi.org/10.1007/BF03172998>.

- Tessier, D., Sarrazin, P., & Ntoumanis, N. (2010). The effect of an intervention to improve newly qualified teachers' interpersonal style, students motivation and psychological need satisfaction in sport-based physical education. *Contemporary Educational Psychology*, 35, 242–253. <http://dx.doi.org/10.1016/j.cedpsych.2010.05.005>.
- Vansteenkiste, M., Simons, J., Soenens, B., & Lens, W. (2004). How to become a persevering exerciser? Providing a clear, future intrinsic goal in an autonomy-supportive way. *Journal of Sport & Exercise Psychology*, 26, 232–249.
- Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: a test of self-determination theory. *Journal of Personality and Social Psychology*, 70, 767–779. <http://dx.doi.org/10.1037/0022-3514.70.4.767>.
- Wolters, C. A. (2004). Advancing achievement goal theory: using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology*, 96, 236–250. <http://dx.doi.org/10.1037/0022-0663.96.2.236>.