Achievement Goal Theory (AGT) has been applied as a core concept for understanding and promoting students’ motivation in physical education (PE) and shows considerable relevance for theoretically and empirically justifying the significance of PE. However, systematically organized reviews of empirical research on AGT are limited to physical activities without explicit PE perspective. First, we aimed to compile basic tenets of AGT and its pedagogical potential for PE. Second, to bring together key findings and discuss future research, we systematically examined the existing empirical literature that applied AGT constructs in both observational and interventional PE settings. We searched the Web of Science, Scopus, Education Source, ERIC, SPORTDiscus, Physical Education Index, PsychInfo, and PsychArticles databases to identify English-language peer-reviewed journal articles with no restriction to publication date. The review was conducted according to the PRISMA Statement. Two independent reviewers screened all studies identified for eligibility, and assessed the methodological quality as well as the risk of bias. A total of 91 studies were included for analysis. Most of the studies (70) were observational, 21 studies were intervention based. On average, the methodological quality of the included studies was moderate and the risk of bias was moderate to high. Mastery goals, mastery-approach goals, and mastery climates appear to be highly relevant for supporting multiple political and curricular PE aims such as psychological well-being, motor skill development, general sports participation, prosocial behavior, and aspects of healthy living. Achievement goal profiles combining high mastery goals, high to low performance goals, and performance-approach goals partly show desirable functions. The results provide comprehensive information for planning and shaping PE
INTRODUCTION

A main purpose of formal educational institutions in differentiated modern societies is to transmit, mediate, and transform cultural heritage and social life (Fend, 2006). In elementary and secondary schooling, an expanding range of linguistic, scientific, technical, artistic, and practical subjects were established to serve this function (Goodson, 1990). Although its status has been embattled from the beginning, physical education (PE) has been a stable constant in international education policy, school curricula, and hour boards since the late nineteenth century (Marshall and Hardman, 2000; Kirk, 2001; Houlihan and Green, 2006; Penney, 2008; Tinning, 2012). In order to legitimate the disputed significance of PE, a dual justification strategy has prevailed in many countries around the world (see Pühse and Gerber, 2005 for international confirmations). On one hand, PE is supposed to provide education about, in, and of the physical. The primary subject matter is therefore bodily movement, play, and sport. PE opens up multifarious movement cultures and provides necessary skills for self-determined lifelong participation in sports and other fields of physical activities. On the other hand, PE is assumed to contribute to education through the physical. Here, bodily movement, games, and sports are considered to be useful instruments for stimulating general educational objectives external to those of the physical. PE can help to support moral development, develop prosocial attitudes and competencies, and prepare for an active and healthy lifestyle. Furthermore, positive effects concerning psychological well-being and emotional growth are frequently mentioned, as well as benefits for cognitive capacity and the promotion of good citizenship (for PE legitimation strategies in general see e.g., Williams, 1930; Arnold, 1979; Carr, 1979; Sallis and McKenzie, 1991; Kirk, 1998; Eldar and Ayvazo, 2009; Siedentorp, 2009; Krüger, 2010).

Regarding this ambitious and at most partially scientifically saturated plethora of aims and benchmarks (for critical empirical summaries of PE claims and realities see Hardman, 2008; Bailey et al., 2009; Bailey, 2018), comprehensive psychological research in school PE settings has stressed the significant meaning and predictive power of students’ motivation. To promote active engagement, optimal success, and goal attainment in school PE, it is crucial for learners to possess motivation to participate, perform, and learn (Biddle, 1999; Chen, 2001; Sun, 2016). Among the multitude of social-psychological approaches to scientifically and practically conceptualize motivation, motivated learning processes, and motivated learning outcomes, Self-Determination Theory (SDT; Ryan and Deci, 2002), and Achievement Goal Theory (AGT; Dweck, 1984; Nicholls, 1984; Ames, 1992b; Elliott, 1999) have successfully proven their educational suitability as core concepts to understand and promote students’ motivation in PE (Lirgg, 2006).

While SDT is the most widely used theoretical framework—integrating theoretical considerations, an extensive body of empirical studies, narrative reviews, and systematic reviews of research (Bryan and Solmon, 2007; Ntoumanis and Standage, 2009; Van den Berghe et al., 2014; Sun et al., 2017), AGT also provides eminent fit for sports, structured physical activities (exercise), and PE settings (Treasurer and Roberts, 1995; Weigand et al., 2001; Roberts, 2012; Chepko and Doan, 2015). Similar to SDT, there are numerous empirical studies and some narrative summaries regarding PE and AGT (Ntoumanis and Biddle, 1999; Duda and Ntoumanis, 2003; Chen and Ennis, 2004; Cecić Erpič, 2011; Rudisill, 2016; Liu et al., 2017). However, systematically organized reviews are limited to physical activities without explicit PE perspectives, e.g., competitive sports, physical activity, or motor skill programming (Biddle et al., 2003; Harwood et al., 2015; Lochbaum et al., 2016; Palmer et al., 2017). As the results of respective published articles are not yet systematically appraised and summarized, the justifying power of AGT regarding the subject status of PE and motivational and pedagogically valuable implications for PE lessons are of limited informative value. The present study builds on this issue, taking two main steps into account: first, we compile basic tenets of AGT to discern its pedagogical potential for PE. Second, we provide a systematic review of empirical studies exploring motivation in PE from an AGT perspective.

FOUNDATIONS OF AGT AND ITS POTENTIAL FOR PE

Achievement motivation can be referred to as the drive for success or the attainment of excellence (Harackiewicz et al., 1997). This type of motivation is especially important in academic settings as it has been found to be a predictor of students’ perceptions of their school environment, school engagement, and academic success (e.g., Busato et al., 2000; Wang and Eccles, 2013). AGT stems from social-cognitive and interactionist approaches of (achievement) motivation which involve the interplay of dispositional, contextual, and behavioral factors (Bandura, 2012). Specifically, AGT encompasses individuals’ achievement goals and their perceived motivational
climates, which in turn leads to achievement goal involvement and resulting achievement behaviors. We have visualized these basic tenets of Achievement Goal Theory in Figure 1.

In AGT, achievement goals are described as cognitive representations of desired (or undesired) outcomes in achievement contexts that an individual seeks to approach or avoid (Hulleman et al., 2010). In earlier research on AGT, two constructs of achievement goals were distinguished: mastery and performance (Ames and Archer, 1988). On a general level, mastery goals describe developing competence through learning, increasing understanding, improving over time, and ultimately mastering a task based on intrapersonal standards. Performance goals on the other hand, describe a striving to demonstrate competence or skill to outperform others based on normative standards (e.g., Nicholls, 1984; Dweck, 1986; Ames and Archer, 1988; Dweck and Leggett, 1988; Elliott and Dweck, 1988).

In the course of theoretical differentiation, the dichotomous framework of AGT has been extended by the approach or avoidance valence (see Elliot and Church, 1997 for the trichotomous model; Elliot and McGregor, 2001 for the 2 × 2 model). This distinction was developed to better understand the differences between mastery and performance goals and resulting outcomes (Murayama et al., 2012) and involves approaching a positive prospect or avoiding a negative prospect. Specifically, mastery-approach goals describe a striving to develop and cultivate competencies while mastery-avoidance goals are centered on avoiding, losing, or not developing competencies. Performance goals however, contain appearance, normative, and (more recently identified) evaluative aspects which an individual can either strive to approach or avoid (Hulleman et al., 2010). The appearance aspect entails a striving to appear competent in front of others (e.g., wanting others to believe one is skilled). The normative aspect concerns a striving to perform better than others based on objective standards (e.g., wanting to perform better on a test than others). The evaluative aspect involves a combination of appearance and normative aspects, describing a striving to demonstrate greater ability as evaluated by a superior figure when compared to the performance of others (e.g., wanting a teacher to believe one is more skilled than others). Each of these aspects have been described and labeled in a variety of different ways throughout the literature (Hulleman et al., 2010). Moreover, it is important to note as described in the work of Senko et al. (2011), that variance exists in the conceptual understandings of performance goals in research. In general, previous literature on the topic of AGT in education shows that mastery-approach goals are consistently linked to adaptive processes and outcomes in students (Hulleman et al., 2010), while findings on mastery-avoidance goals remain inconclusive. Results concerning performance-approach goals also appear to be rather inconclusive, potentially due to the lack of previous differentiation between appearance, normative, and evaluative aspects, while performance-avoidance goals have been clearly linked to maladaptive outcomes (Harackiewicz et al., 2002; Hulleman et al., 2010).

In AGT, it is assumed that motivational climate, or classroom goal structures, affect individual outcomes through their individual achievement goals (which are set in accordance to the surrounding goal climate). Thus, it is not only an individual’s internal attributes that affect their goal setting processes, but also the specific contextual circumstances in which the achievement task is defined (Ames, 1992a). Similar to achievement goals, there are two types of motivational climate: mastery goal structure, with a focus on intrapersonal standards for the individuals involved in the environment, and performance goal structure, with a focus on comparison with others and demonstrating superiority. Motivational climate can be influenced by the attitudes and behaviors of persons involved in the particular sport or class (e.g., peers, parents, coaches, teachers, etc.). Epstein (1988, 1989), Ames and Archer (1988), and Ames (1992a) provided a well-known theoretical framework used by researchers known as TARGET for identifying classroom dimensions that influence motivation and determining whether a classroom climate is mastery or performance oriented. This system entails looking at instructional practices of the teacher related to: T (task assignments), A (authority relations), R (recognition systems), G (grouping procedures), E (evaluation practices), and T (use of time). The aspects involved in TARGET can be assessed and altered by teachers to create a more positive motivational climate. Again, there is a consensus within the literature that mastery climates (in comparison to performance climates) are beneficial for PE students (see Braithwaite et al., 2011 for a PE summary based on TARGET strategies).

In sum, the logic behind mastery and mastery-approach goals and mastery climate is of pedagogical significance for PE. PE lessons continually evoke salient and publicly expressed achievement situations, i.e., physical demonstrations of abilities, standards of excellence, and performance evaluations about, in, for, and through bodily movements, play, and sports. Within mastery goals, PE students thereby focus on effort, personal improvement, and task mastery. These aspects have commonly been associated with positive education outcomes. Students who are highly mastery-oriented are likely to enjoy the process of moving and learning, to select more challenging tasks, and to help others. They tend to perceive success and failure in association with exertion, to persist in practice in the face of difficulty or failure, and to show enhanced moving and learning benefits. The presence of performance and performance-avoidance goals on the contrary is typically assumed to show opposite or no such consequences. A mastery climate entails collaborative tasks, democratic leadership, recognition for effort and improvement, private and individual evaluation, and sufficient time for everyone to learn. In line with mastery goals, the possibility to structure educational environments is linked with enhanced motivation, adaptive patterns of behavior, and desirable curricular effects. Conversely, students who perceive a performance-oriented climate are deemed at risk of avoiding challenges and of being unwilling to expend effort, especially when they experience failure or encounter difficulty performing.

1Looking at the relevant literature, there is no uniform terminology. The same constructs have also been labeled as goals of learning and performance (Dweck, 1984), task and ability (Maehr and Midgley, 1991), and task- and ego-involved (Nicholls, 1984). According to clarity and consistency, this article refers to the terms mastery and performance.
Jaitner et al. Achievement Goal Motivation in PE

FIGURE 1 | Basic tenets of Achievement Goal Theory.

METHODS

The systematic review was conducted according to the PRISMA Statement (Moher et al., 2015).

Search Strategy and Eligibility Criteria

In order to identify all relevant articles, a systematic literature search was conducted. The search string combined two key elements: objective and setting. Terms of objective followed the social-cognitive and interactionist structure of AGT and comprised aspects of achievement goals and motivational climate: [achievement goal theory] OR [achievement goal OR task goal OR task-involved goal OR ego-involved goal OR performance goal OR learning goal OR ability goal] OR [motivational climate OR performance (motivational) climate OR mastery (motivational) climate OR goal structures]. For setting, the single term physical education was fixed without meaningful alternatives. The search was performed on 6 April 2018 using four thematic groups of electronic databases: general databases (Web of Science, Scopus), educational databases (Education Source, ERIC), sport-scientific databases (SPORTDiscus, Physical Education Index), and psychological databases (PsychInfo, PsychArticles). Additionally, reference lists and citations of included articles were screened to identify additional studies. The following eligibility criteria were compulsory in order to be included in the review: (1) original empirical research published in English language peer-reviewed journals, (2) relationship with at least one term of every term type, (3) application of a valid quantitative assessment of achievement goal motivation, (4) genuine reference to physical education (not e.g., to extracurricular school sports, physical activity, sports in general, competitive sports, or high-performance sports); studies that use physical education lessons as a setting for proband acquisition but predict non-educational goals were not considered, (5) adequate methodological quality. Studies involving special educational conditions were excluded from the review. No restrictions on publication periods were given due to the relatively recent origin age of AGT. All quantitative study designs were allowed.

Study Selection, Data Extraction, and Analysis

Two independent reviewers (DJ, RR) conducted and carefully documented a stepwise literature search. First, the search results were exported into EndNote X8 reference management software and duplicates were removed. Second, the titles, abstracts, and full-texts of the exported studies were gradually screened for eligibility. Based on the given information within the titles and abstracts, the reviewers made decisions about inclusion or exclusion. If the abstract indicated that the study fulfilled the eligibility criteria, the abstract was missing, or the abstract did not provide sufficient information for selection decision, both reviewers assessed the full-texts of the articles for eligibility. Third, all reference lists and citations of included articles listed in Scopus were reviewed using the same procedure to identify additional studies. At each step, any discrepancies regarding criteria fulfillment were consensually resolved by discussion. Fourth, methodological quality was assessed independently by two reviewers (RR, JB) to exclude studies without adequate fit. The extracted target data included: (a) study characteristics (author, date of publication, journal, objective of the study, and...
research design), (b) sample characteristics (mean age, sample size, gender distribution, country, and school type), (c) study design (temporal setup, applied instruments of motivational assessment, dependent variables, methodology, and analytic process), (d) main outcomes of the studies, and (e) barriers and limitations. When essential information was missing from the full-texts, the corresponding authors were contacted to obtain the missing details. The search and selection processes are illustrated in a flow chart. The main descriptive characteristics of the studies were outlined using a detailed table. The relevant results of the included studies were evaluated via narrative synthesis.

Critical Appraisal
Each article meeting the eligibility criteria 1–4 was critically appraised to judge the methodological quality and to determine the extent to which a study excluded or minimized the possibility of bias in design, conduct, and analysis. (a) According to the encountered study designs, the methodological quality was evaluated using the Joanna Briggs Institute (JBI) checklists for randomized controlled trials (RCT), non-randomized experimental studies (NRES), cohort studies (LS), and analytical cross-sectional studies (CS), rating the scope of methodological quality over eight to 13 domains (Joanna Briggs Institute [JBI], 2018). All corresponding articles were independently assessed by two reviewers (RR, JB) and were then discussed until consensus was achieved. Cohen’s kappa was calculated as a measure of initial inter-observer agreement. The critical appraisal had a dual function in the review. First, it served to provide information about the technical quality and thus, the degree of validity of the included studies. Second, it was used to exclude approaches with limited methodological quality. The preliminarily included studies were therefore ranked based on the results of the methodological quality assessment. Given the different numbers of items in the applied appraisal tools, the comparison of the studies required a statistical procedure that combined qualitative and quantitative elements: the ranking was conducted using pairwise comparison, a recognized method applicable to transfer a given number of different entities into a scale of relative importance (Bramley and Oates, 2011). In the current study, the included studies were compared and it was evaluated whether the methodological quality of any selected study was better (+1), equal (0), or worse (−1) than the comparative approaches. The total of the comparisons then resulted in a ranking with which relatively inferior studies could be identified and excluded. (b) The risk of bias assessment followed the schedule by Porritt et al. (2014) and referred to potential sources of bias in selection, performance, detection, and attrition.

RESULTS
The initial literature search resulted in 1787 hits. Nine hundred and sixty-two remained after deleting the duplicates. After screening the titles and abstracts, 139 studies were included in the full-text audit. One hundred and thirteen studies met the eligibility criteria 1–4. The reference search added three studies, and the cited-by-search added one study to be included to the sample in process. After the methodological quality assessment of all preliminarily included studies, 91 studies were finally chosen for the systematic review (see Figure 2).

Methodological Quality and Risk of Bias
Initial inter-observer agreement on all assessed studies was weak to moderate for LSs ($\chi^2 = 0.58$), moderate for RCTs ($\chi^2 = 0.67$), and NRESs ($\chi^2 = 0.77$), and strong for CSs ($\chi^2 = 0.90$; McHugh, 2012). For all deviations, complete agreement was reached by discussion. Most of the included studies were CSs ($n = 65$; 71.4%), followed by NRESs ($n = 15$; 16.5%), RCTs ($n = 6$; 6.6%), and LSs ($n = 5$; 5.5%). The overall level of evidence therefore can be considered relatively low with some upward outliers (Joanna Briggs Institute [JBI], 2014). The final results of the methodological quality evaluation for the included studies are presented at length in the Supplementary Tables 1–4. On average, the results provide a rather middle-rate methodological quality impression. An exception is the NRESs. Here, the methodological quality can be considered comparatively high.

The risk of bias assessment indicated some ambiguities that conceivably restricted the validity of the appraised studies. All RCTs reported a random generation of groups. The method of the random sequence generation, however, remained indeterminate. In combination with constantly unclear allocation concealment, a selection bias might have occurred. This also applies to the included NRESs, LSs, and CSs. The NRES were conducted in a school setting and meet established structures there. This means that true randomization was excluded or possible with increased effort only. Instead, the allocation to the experimental groups took place in recourse to existing school PE groups and classes. Selection bias in observational studies particularly occurs when the subjects studied are not representative of the target population about which conclusions are to be drawn: here, the included LSs and CSs almost always introduced a multitude of demographic variables and structural data. However, information on the background of the sample design and the intended range of the studies was largely missing. Contrary to the thorough blinding of the participants, the educational character of the intervention studies may have potentially obstructed the double-blinded study designs. In both RCTs and NRESs, the interventions each required competence for and experience in the instruction and implementation of specific motivational climates. The interventions were therefore carried out by the expert researchers or by well-trained and supervised physical educators. Thus, blinding of personnel was ruled out and those delivering treatment were given the opportunity to behave differently with the participants from the different groups. Therefore, an unclear risk of performance bias can be assumed. Only one RCT was appraised as having a low risk of detection bias with regard to blinding of outcome assessors (Logan et al., 2014). All other included studies did not provide sufficient information on outcome assessment but at least partially mitigated a possible distortion by usually using valid instruments and questionnaires. Potential attrition bias results from incomplete follow-up or losses in follow-up that are inadequately taken into account: six NRESs (Morgan and Carpenter, 2002; Jaakkola and Liukkonen, 2006; Viciana et al.,
FIGURE 2 | Flow chart of study search and selection process.
TABLE 1 | Descriptive characteristics of the included observational studies.

<table>
<thead>
<tr>
<th>Source</th>
<th>N (girls)</th>
<th>Mean age + SD</th>
<th>Nation</th>
<th>School type</th>
<th>Study design</th>
<th>AGT frame</th>
<th>AGT climate</th>
<th>AGT goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agbuga and Xiang, 2008</td>
<td>229 (107)</td>
<td>8th 14.05 ± 0.67 11th 17.28 ± 0.90</td>
<td>Turkey</td>
<td>Sec</td>
<td>cs</td>
<td>Tricho</td>
<td>–</td>
<td>TEOSQ/AGQ</td>
</tr>
<tr>
<td>Agbuga, 2010</td>
<td>229 (107)</td>
<td>8th 14.05 ± 0.67 11th 17.28 ± 0.90</td>
<td>Turkey</td>
<td>Sec</td>
<td>cs</td>
<td>Tricho</td>
<td>–</td>
<td>TEOSQ/AGQ</td>
</tr>
<tr>
<td>Agbuga, 2014</td>
<td>229 (107)</td>
<td>8th 14.05 ± 0.67 11th 17.28 ± 0.90</td>
<td>Turkey</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ</td>
<td>–</td>
</tr>
<tr>
<td>Baena-Extremera et al., 2015a</td>
<td>758 (411)</td>
<td>15.22 ± 1.27</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Baena-Extremera et al., 2015b</td>
<td>2002 (1032)</td>
<td>14.99 ± 1.43</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Baena-Extremera et al., 2013</td>
<td>2002 (1032)</td>
<td>14.99 ± 1.43</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Baena-Extremera et al., 2016</td>
<td>2168 (1037)</td>
<td>12.49 ± 0.81</td>
<td>Spain, Costa Rica, Mexico</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>2 × 2 AGQ-PE</td>
<td></td>
</tr>
<tr>
<td>Bakirtzoglou and Ioannou, 2010</td>
<td>200 (100)</td>
<td>15.42 ± 1.23</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Baric and Cecić Erpić, 2014</td>
<td>594 (311)</td>
<td>11–18</td>
<td>Croatia</td>
<td>Elem/sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Barkoukis et al., 2012</td>
<td>368 (172)</td>
<td>14.83 ± 0.96</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PSGS, LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Bortoli et al., 2014</td>
<td>167 (85)</td>
<td>14.51 ± 0.52</td>
<td>Italy</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>TMOPC</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Bryan and Solmon, 2012</td>
<td>114 (57)</td>
<td>Not reported</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Camacho et al., 2008</td>
<td>938 (485)</td>
<td>14.8 ± 0.91</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ-2 POSQ</td>
<td>–</td>
</tr>
<tr>
<td>Carr and Weigand, 2001</td>
<td>266 (121)</td>
<td><em>♂</em> 12.4 ± 0.92 <em>♀</em> 13.0 ± 0.89</td>
<td>UK</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PECCS</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Carr and Weigand, 2002</td>
<td>266 (121)</td>
<td><em>♂</em> 12.4 ± 0.92 <em>♀</em> 13.0 ± 0.89</td>
<td>UK</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PECCS</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Carr, 2006</td>
<td>193 (92)</td>
<td>t1 11 ± 2 mo.</td>
<td>UK</td>
<td>Sec</td>
<td>cs/l/s</td>
<td>Tricho</td>
<td>PMCSQ-2 PALS</td>
<td></td>
</tr>
<tr>
<td>Cecchini Estrada et al., 2011</td>
<td>395 (209)</td>
<td>15.68 ± 1.06</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>2 × 2 AGQ-PE</td>
<td></td>
</tr>
<tr>
<td>Cox and Williams, 2008</td>
<td>518 (289)</td>
<td>11.38 ± 0.68</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>dicho</td>
<td>PMCSQ</td>
<td>–</td>
</tr>
<tr>
<td>Curry et al., 2002</td>
<td>682 (0)</td>
<td>14.03 ± 0.70</td>
<td>France</td>
<td>Sec</td>
<td>cs</td>
<td>Tricho</td>
<td>–</td>
<td>AGQ</td>
</tr>
<tr>
<td>Digelidis and Papaioannou, 1999</td>
<td>674 (355)</td>
<td>10–17</td>
<td>Greece</td>
<td>Elem/sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Dorobantu and Biddle, 1997</td>
<td>145 (84)</td>
<td>15 ± 2 mo.</td>
<td>Romania</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PECCS</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Fernández-Rio et al., 2014</td>
<td>507 (240)</td>
<td>14.37 ± 1.69</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ-2</td>
<td>–</td>
</tr>
<tr>
<td>Fernández-Rio et al., 2012</td>
<td>304 (177)</td>
<td>15.6 ± 1.04</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>2 × 2 AGQ-PE</td>
<td></td>
</tr>
<tr>
<td>Flores et al., 2008</td>
<td>2993 (1615)</td>
<td>13.5 ± 1.7</td>
<td>Colombia</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Gao et al., 2011</td>
<td>194 (101)</td>
<td>12.4 ± 1.0</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>2 × 2 AGQ-S</td>
<td></td>
</tr>
<tr>
<td>Garn et al., 2011</td>
<td>105 (48)</td>
<td>15.16 ± 0.66</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>2 × 2 AGQ-PE</td>
<td></td>
</tr>
<tr>
<td>Gómez-López et al., 2015</td>
<td>846 (363)</td>
<td>15.47 ± 1.32</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ-2 POSQ</td>
<td>–</td>
</tr>
<tr>
<td>Goudas and Biddle, 1994</td>
<td>254 (100)</td>
<td>13–15</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECO</td>
<td>–</td>
</tr>
<tr>
<td>Guan et al., 2006</td>
<td>544 (327)</td>
<td>16.43 ± 1.23</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>2 × 2 AGQ-PE</td>
<td></td>
</tr>
<tr>
<td>Halvari et al., 2011</td>
<td>t1 172</td>
<td>t2 152</td>
<td>Norway</td>
<td>Sec</td>
<td>ls</td>
<td>Tricho</td>
<td>MCSQ</td>
<td>AGQ</td>
</tr>
<tr>
<td>Hsu et al., 2017</td>
<td>st1 287 (132)</td>
<td>st1 14.2</td>
<td>Taiwan</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>–</td>
<td>AGPEQ</td>
</tr>
<tr>
<td></td>
<td>st2 296 (132)</td>
<td>st2 14.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Source</th>
<th>N (girls)</th>
<th>Mean age + SD</th>
<th>Nation</th>
<th>School type</th>
<th>Study design</th>
<th>AGT frame</th>
<th>AGT climate</th>
<th>AGT goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jaakkola et al., 2015</td>
<td>4397 (2594)</td>
<td>14–15</td>
<td>Finland</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>MCPEQ</td>
<td>–</td>
</tr>
<tr>
<td>Johnson et al., 2017</td>
<td>290 (150)</td>
<td>11–14</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ-2</td>
<td>–</td>
</tr>
<tr>
<td>Kalaja et al., 2009</td>
<td>370 (189)</td>
<td>13.08 ± 0.25</td>
<td>Finland</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>IMCPEQ</td>
<td>–</td>
</tr>
<tr>
<td>Kavussanu, 2007</td>
<td>156 (75)</td>
<td>14.85 ± 1.26</td>
<td>UK</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>POSQ</td>
</tr>
<tr>
<td>Kokkonen et al., 2013</td>
<td>t1 258 (122)</td>
<td>15–16</td>
<td>Finland</td>
<td>Sec</td>
<td>Is</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>POSQ</td>
</tr>
<tr>
<td>Kouli and Papaioannou, 2009</td>
<td>1305 (683)</td>
<td>14.5 ± 0.49</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Liukkonen et al., 2010</td>
<td>338 (163)</td>
<td>11–12</td>
<td>Finland</td>
<td>Elem</td>
<td>cs</td>
<td>Dicho</td>
<td>MCPEQ</td>
<td>–</td>
</tr>
<tr>
<td>Moreno-Murcia et al., 2011</td>
<td>566 (306)</td>
<td>14.5 ± 0.49</td>
<td>Spain</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ-2</td>
<td>POSQ</td>
</tr>
<tr>
<td>Mouratidis et al., 2010</td>
<td>st1: 372 (199)</td>
<td>st1 11.39 ± 0.65</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>Not clear</td>
</tr>
<tr>
<td></td>
<td>st2: 339 (153)</td>
<td>st2 13.97 ± 0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouratidis et al., 2009</td>
<td>319 (165)</td>
<td>Not reported</td>
<td>Greece</td>
<td>Elem</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Ommundsen and Eikanger Kval, 2007</td>
<td>194 (94)</td>
<td>16</td>
<td>Norway</td>
<td>Elem/Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PECCS</td>
<td>–</td>
</tr>
<tr>
<td>Ommundsen, 2001</td>
<td>343 (177)</td>
<td>15–16</td>
<td>Norway</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>POSQ</td>
</tr>
<tr>
<td>Ommundsen, 2004</td>
<td>273 (148)</td>
<td>15.3 ± 0.70</td>
<td>Norway</td>
<td>Sec</td>
<td>cs</td>
<td>Tricho</td>
<td>–</td>
<td>GOS-A</td>
</tr>
<tr>
<td>Ommundsen, 2006</td>
<td>273 (148)</td>
<td>15.3 ± 0.70</td>
<td>Norway</td>
<td>Sec</td>
<td>cs</td>
<td>Tricho</td>
<td>PECCS</td>
<td>GOS-A</td>
</tr>
<tr>
<td>Papaioannou and Macdonald, 1993</td>
<td>211 (118)</td>
<td>14–17</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Papaioannou, 1997</td>
<td>1393</td>
<td>12–18 y.</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>–</td>
</tr>
<tr>
<td>Papaioannou, 1998</td>
<td>674 (355)</td>
<td>10–17 y.</td>
<td>US</td>
<td>Elem/Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Parish and Treasure, 2003</td>
<td>442 (229)</td>
<td>12.56 ± 0.96</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Parisi et al., 2015</td>
<td>428 (247)</td>
<td>13.3 ± 0.97</td>
<td>Greece</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>PALS</td>
</tr>
<tr>
<td>Ruiz-Juan et al., 2018</td>
<td>2089 (1037)</td>
<td>12.49 ± 0.81</td>
<td>Spain</td>
<td>Sec</td>
<td>Tricho</td>
<td>–</td>
<td>PTEGQ</td>
<td>–</td>
</tr>
<tr>
<td>Shen et al., 2007</td>
<td>202 (99)</td>
<td>11–13 y.</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Tricho</td>
<td>–</td>
<td>AGQ</td>
</tr>
<tr>
<td>Solmon, 2006</td>
<td>278 (140)</td>
<td>Students</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Spray, 2002</td>
<td>488 (256)</td>
<td>14.18 ± 0.49</td>
<td>UK</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>–</td>
</tr>
<tr>
<td>Sproule et al., 2007</td>
<td>st1 230 (177)</td>
<td>st1 15.2 ± 0.80</td>
<td>Singapore</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>TEOSQ</td>
</tr>
<tr>
<td></td>
<td>st2 802 (402)</td>
<td>st2 13.4 ± 0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standage et al., 2003</td>
<td>296 (138)</td>
<td>13.6 ± 0.60</td>
<td>UK</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>PMCSQ</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Theodosiou et al., 2008</td>
<td>494 (277)</td>
<td>Not reported</td>
<td>Greece</td>
<td>Elem/Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>LAPOPECQ</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Treasure, 1997</td>
<td>233 (119)</td>
<td>10–12</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>CAGQ</td>
</tr>
<tr>
<td>Tzetzis et al., 2002</td>
<td>112 (62)</td>
<td>9–12</td>
<td>US</td>
<td>Elem</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Walling and Duda, 1995</td>
<td>144 (78)</td>
<td>15.2 ± 0.78</td>
<td>US</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Wang and Liu, 2007</td>
<td>343 (343)</td>
<td>13.51 ± 0.37</td>
<td>Singapore</td>
<td>Sec</td>
<td>cs</td>
<td>Dicho</td>
<td>–</td>
<td>TEOSQ</td>
</tr>
<tr>
<td>Wang et al., 2010</td>
<td>782 (524)</td>
<td>15.24 ± 1.45</td>
<td>Singapore</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>LAPOPECQ</td>
<td>AGPEQ</td>
</tr>
<tr>
<td>Wang et al., 2008</td>
<td>484 (262)</td>
<td>14.32 ± 0.98</td>
<td>Singapore</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>TMCPQ</td>
<td>AGPEQ</td>
</tr>
<tr>
<td>Wang et al., 2016</td>
<td>1810 (1145)</td>
<td>14–19</td>
<td>Singapore</td>
<td>Sec</td>
<td>cs</td>
<td>2 × 2</td>
<td>LAPOPECQ</td>
<td>AGPEQ</td>
</tr>
<tr>
<td>Warburton and Spray, 2013</td>
<td>403 (203)</td>
<td>11.13.16 ± 0.86</td>
<td>UK</td>
<td>Sec</td>
<td>Is</td>
<td>2 × 2</td>
<td>–</td>
<td>AGQ-S</td>
</tr>
</tbody>
</table>

(Continued)
Continued

Table 1 provides a summary of the descriptive characteristics across the observational studies, including 74 independent samples (70 studies). The dates of publication ranged from 1993 to 2018, with a clear focus on the last 10 years. Most of the studies were conducted in individual European countries and the USA. Two studies were cross-nationally configured (Baena-Extremera et al., 2016; Ruiz-Juan et al., 2018). In total, 42,004 participants were included in the 69 CS samples, at least 21,310 were female (about 52.5%; one study made no specification about gender). One thousand three hundred ninety-one participants attended all measurement time points of the five LSs samples (about 55.9% female; two studies did not report gender distribution). The average sample size for CSs and LSs was 609 and 278, respectively. In the majority of the cases, the samples consisted of secondary education students. Six trials each included participants from elementary schools or between school types. The studies were mostly based in the dichotomous framework of AGT (n = 49; 70%), with the trichotomous framework (n = 9; 12.9%) and the 2 × 2 achievement goal framework (n = 12; 17.1%) being selectively employed. Perceived achievement goals were primarily obtained by the TEOSQ (n = 24; 47.1%), the AGQ (n = 7; 13.7%), the 2 × 2 AGQ (n = 6; 11.8%), and the POSQ (n = 6; 11.8%). To assess the perceived motivational climate, the LAPOPEQCQ (n = 18; 40.9%), the PMCSQ-2 (n = 6; 13.6%), the PMSC (n = 6; 13.6%), and the PECCS (n = 5; 11.4%) were mainly used.

Supplementary Table 5 compiles the analytical approaches and main outcomes of the included observational studies individually. The comprising narration of the results is presented below using four main categories (see Figure 3). The categories were inductively derived from the objectives and the structure of the included studies by the first author, an experienced qualitative analyst, and thus provide an ordering solution closely inherent to the data material.

1. A first group of studies examined group differences (n = 15) of achievement goals and perceived motivational climates in students, namely sociodemographic differences (e.g., gender, age, country, and school):

- Concerning gender, seven studies (Walling and Duda, 1995; Digelidis and Papaioannou, 1999; Carr and Weigand, 2001; Flores et al., 2008; Cecchini Estrada et al., 2011; Moreno-Murcia et al., 2011; Baric and Ceci’c Erpić, 2014) found significant differences, with females scoring higher in mastery and mastery-avoidance goals (Digelidis and Papaioannou, 1999; Cecchini Estrada et al., 2011; 2007; Erturan-lker and Demirhan, 2013; Abós et al., 2016; Bortoli et al., 2017) and three LSs (Braithwaite et al., 2011; Halvari et al., 2011; Warburton and Spray, 2013) showed appropriate treatment of drop-outs. For all RCTs and the majority of the NRESs, drop-outs of participants remained unclear, thus leaving room for biasing differences between initial and ending samples. In summary, the risk of bias assessment revealed an enormous range of methodological issues and indicated moderate to high risk of bias within the appraised approaches.
Baric and Cecić Erpić, 2014), and males scoring higher in performance goals, performance-approach goals, and performance climates (Walling and Duda, 1995; Carr and Weigand, 2001; Flores et al., 2008; Cecchini Estrada et al., 2011; Moreno-Murcia et al., 2011). Conversely, some studies observed no significant gender differences (Tzetzis et al., 2002; Agbuga, 2010).

- Regarding age differences, the majority of the studies reported significant effects (Digelidis and Papaioannou, 1999; Xiang and Lee, 2002; Agbuga and Xiang, 2008; Theodosiou et al., 2008; Bryan and Solmon, 2012; Baric and Cecić Erpić, 2014). However, the results varied, with some studies finding older students (11th grade) to be more inclined toward performance goals and performance climates than younger students (elementary school-aged or 4th/8th grade) (Xiang and Lee, 2002; Theodosiou et al., 2008). The opposite was also found for performance-approach goals (Agbuga and Xiang, 2008). A handful of studies established older students (such as 8th grade or senior high school-aged) to have lower perceptions of mastery climate than their younger counterparts (such as elementary school-aged or 6th/7th grade) (Digelidis and Papaioannou, 1999; Flores et al., 2008; Bryan and Solmon, 2012). On the contrary, it was also found that older students were more mastery and performance-oriented (Baric and Cecić Erpić, 2014).

- A single study (Ruiz-Juan et al., 2018) investigated differences in perceived motivational climates of students depending on the country in which they went to school, and showed that regardless of country (Costa Rica, Mexico, Spain), mastery climates remained higher than performance-approach and performance-avoidance climates. One study found that schools differed in students’ perceptions of motivational climate (Johnson et al., 2017).

2. A second group of studies focused on interrelations (n = 16), that is, on undirected connections between constructs of AGT or relations between variables of AGT and other psychological variables:

- When addressing the interrelations between constructs of AGT, it was found that mastery goals were positively related to mastery climate, while performance goals were positively related to performance climates (Carr and Weigand, 2001, 2002; Sproule et al., 2007). Goal profiles high in mastery-approach and mastery-avoidance goals showed the highest scores for mastery and performance climates (Wang et al., 2008). Mastery climate was positively related to performance climate (Sproule et al., 2007). Climate profiles with consistent high mastery and low performance climates were related to high levels of mastery goals and decreased levels of performance-avoidance goals. Climate profiles with consistent low mastery and high performance climates were related to decreased mastery goal scores and increased performance-avoidance scores (Carr, 2006).

- Concerning relations with other psychological variables, mastery goals were positively related to enjoyment and perceived competence (Baric and Cecić Erpić, 2014), cooperation and self-esteem (Papaioannou and Macdonald, 1993), out-of-school sport (Papaioannou, 1997), social goals (Solmon, 2006), participation in vigorous physical activity (Tzetzis et al., 2002), and intent for participation in PE (Wang et al., 2016). Mastery climate was positively related to more autonomous levels of self-determined motivation (Parish and Treasure, 2003; Sproule et al., 2007, 2012), positive attitudes (Wang et al., 2008; Bryan and Solmon, 2012), enjoyment (Jaakkola et al., 2015), and perceived competence and intention (Sproule et al., 2007; Wang et al., 2008). Performance climate was positively related to lower levels of enjoyment (Jaakkola et al., 2015), external regulation and amotivation (Parish and Treasure, 2003), and perceived competence (Sproule et al., 2007). Compared to goal profiles high in performance goals and low in mastery climates.
goals, goal profiles high in mastery goals were related to longer participation in vigorous physical activity regardless of the students’ performance goals (Tzetzis et al., 2002). In addition, high mastery goal profiles were positively related to higher perceived global flow state (Camacho et al., 2008), more enjoyment, intention, perceived competence, and higher belief purposes (Wang et al., 2008), while individuals in both low mastery and low performance clusters had lower scores for success being intrinsic and for perceived teacher encouragement (Walling and Duda, 1995).

3. A third group of included studies addressed various predictor variables of perceived PE achievement goals and motivational climates (n = 9). Seven studies thereby explored the predictive utility of a variety of personal predictor variables:

- Three studies assessed the relationships between sport ability beliefs and students’ PE achievement goals (Cury et al., 2002; Wang and Liu, 2007; Warburton and Spray, 2013). Incremental beliefs positively predicted perceived mastery goals and change in mastery-approach goal adoption over time, and negatively predicted performance-approach and performance avoidance goals. Entity beliefs remained without consistent pattern of goal adoption and served as positive predictors of performance goals, performance-approach goals, and performance avoidance goals.

- Two studies each examined the predictive values of social goals, causal attributions, variables of perceived locus of causality (Baena-Extremera et al., 2016; Ruiz-Juan et al., 2018), and perceptions of PE competence (Cury et al., 2002; Warburton and Spray, 2013). To this end, high scores on responsibility goals, internal explanations for events, and perceived competence affirmed the probability of perceived mastery goals, mastery-approach goals, and mastery climates. Perceived performance-avoidance goals were positively predicted by responsibility goals, both attribution patterns, and low perceptions of competence. On the other hand, they were negatively predicted by high perceptions of competence. High scores in internal and external causal attributions and high perceived competence values showed predictive efficiency for performance-approach goals. Regarding the locus of causality, perceived mastery-oriented goals were positively associated with rather beneficial levels of the self-determination continuum, with performance-avoidance goals and climates additionally being predicted by amotivation. For perceived performance-oriented goals, amotivation tended to show the greatest predictive utility, being accompanied by high scores in external regulation for performance-approach and introjected regulation for performance-avoidance goals.

- One study investigated the predictive power of different forms of self-determined motivation (Mouratidis et al., 2010), with autonomously motivated students being more likely to endorse learning, mastery-approach, and ability goals, and less likely to support outcome and normative performance goals than students motivated to engage in PE activities by way of control. In one study, extracurricular sport activities predicted mastery-approach and performance-avoidance goals (Fernández-Rio et al., 2012).

Three studies concentrated on contextual predictor variables of PE students’ achievement goals:

- One study (Cury et al., 2002) examined the predictive power of the perceived motivational climate in the PE classroom and the adoption of achievement goals. Here, perceived mastery climate served as a positive prerequisite of mastery goals. Perceived performance climate proved to be an adaptive contextual condition of both performance-approach and performance-avoidance goals, and showed a negative association with mastery goal adoption.

- Two studies (Carr and Weigand, 2002; Warburton, 2017) explored the relative influences of different social agents on students’ PE achievement goals indicating that for mastery goals, perceived mastery climates emphasized by teachers were the most important positive predictor, followed by perceptions of mastery climates from peers and sporting heroes. For performance goals, perceived performance climates emphasized by peers appeared to be the primary positive variable, ensued by perceptions of performance goals in sporting heroes.

4. A fourth group of included studies focused on PE achievement goals and motivational climates as predictors of different outcome variables (n = 47). Forty-five studies thereby examined the predictive power on personal outcome variables:

- Three studies examined relationships between AGT constructs and identified a consistent reciprocal relationship: performance, performance-approach, and performance-avoidance goals were explained by performance climates; mastery, mastery-approach, and mastery-avoidance goals were explained by mastery climates (Sproule et al., 2007; Wang et al., 2010; Halvari et al., 2011).

- Fifteen studies examined the effects of achievement goals and motivational climates on SDT variables. Mastery goals and mastery climates positively (Kalaja et al., 2009; Halvari et al., 2011; Fernández-Rio et al., 2014) and performance-avoidance goals negatively (Ommundsen, 2004) predicted basic psychological needs (BPN). Regarding the different types of motivational regulation processes established by SDT, mastery goals, mastery-approach goals, and mastery climates showed a positive predictive role for higher levels of self-determination (Goudas and Biddle, 1994; Dorobantu and Biddle, 1997; Parish and Treasure, 2003; Standage et al., 2003; Ommundsen and Eikanger Kval, 2007; Cox and Williams, 2008; Cecchini Estrada et al., 2011; Halvari et al., 2011; Bortoli et al., 2014; Fernández-Rio et al., 2014; Baena-Extremera et al., 2015a; Gómez-López et al., 2015) and a negative predictive role for rather lower levels of self-determination (Ommundsen and Eikanger Kval, 2007; Cecchini Estrada et al., 2011; Halvari et al., 2011; Fernández-Rio et al., 2014). The opposite
was found for performance goals, performance-avoidance goals, and performance climates (Parish and Treasure, 2003; Standage et al., 2003; Ommundsen and Eikanger Kval, 2007; Cecchinì Estrada et al., 2011; Bortoli et al., 2014; Baena-Extremera et al., 2015a), which were especially important in predicting amotivation.

Nine studies examined affective consequences. Mastery goals and mastery climates were positively related to enjoyment (Dorobantu and Biddle, 1997; Ommundsen and Eikanger Kval, 2007; Liukkonen et al., 2010; Wang et al., 2010; Barkoukis et al., 2012; Fernández-Rio et al., 2014), feelings of exertion (Halvari et al., 2011), activating emotions (Mouratidis et al., 2009), and satisfaction (Treasure, 1997), and negatively related to boredom, pressure/tension (Fernández-Rio et al., 2014), and negative emotions (Mouratidis et al., 2009). Performance goals and climates positively predicted boredom (Treasure, 1997; Barkoukis et al., 2012), exertion (Halvari et al., 2011), state anxiety (Liukkonen et al., 2010), and pride (Mouratidis et al., 2009), and negatively predicted enjoyment (Liukkonen et al., 2010; Wang et al., 2010; Barkoukis et al., 2012).

One study showed a positive prognostic relationship between both achievement goals and climates, and dispositional flow. For mastery goals and mastery performance climates, the aforementioned relationships occurred regardless of gender, though with different regression weights (Bakirtzoglou and Ioannou, 2010).

Eleven studies examined behavioral outcomes. With regard to self-reported variables, a positive predictive utility of mastery goals and mastery climates was found for PE attendance (Halvari et al., 2011), disciplined behavior (Moreno-Murcia et al., 2011), moral competences (Parisi et al., 2015), and mastery behaviors (Xiang and Lee, 2002), and a negative predictive utility of both constructs for undisciplined behavior (Moreno-Murcia et al., 2011). In addition, mastery-approach goals predicted three out of five misbehaviors occurring in PE classes: failure to follow directions, poor self-management, and low engagement (Hsu et al., 2017). A positive relationship was found between performance climates and undisciplined behavior (Moreno-Murcia et al., 2011). Performance-approach goals predicted two out of five misbehaviors (aggressive behavior, distracting behavior), and performance-avoidance goals predicted all queried misbehaviors (Hsu et al., 2017). For other-observed behavioral outcomes, in-class physical activities were positively explained by mastery-approach goals (Gao et al., 2011), mastery climates (Parish and Treasure, 2003; Gao et al., 2011) and performance-avoidance goals (Shen et al., 2007). A positive relationship was found between mastery climate and fundamental movement skills (Kalaja et al., 2009). Mastery goals (Xiang et al., 2004a, 2007) and performance goals (Xiang et al., 2004b) positively predicted running performance. Performance climate positively predicted manipulative movement skills (Kalaja et al., 2009).

Ten studies indicated that students' mastery goals (Agbuga and Xiang, 2008; Agbuga, 2010; Baena-Extremera et al., 2013, 2016), mastery-approach goals (Guan et al., 2006; Garn et al., 2011), mastery climates (Liukkonen et al., 2010; Agbuga, 2014; Erturan-Ilker, 2014), performance-approach goals (Guan et al., 2006; Agbuga and Xiang, 2008; Agbuga, 2010), and performance climates (Liukkonen et al., 2010; Agbuga, 2014; Erturan-Ilker, 2014) positively predicted students' perceived effort and/or persistence in PE lessons.

Four studies evaluated relationships between AGT constructs and self-regulation practices. To this end, metacognitive strategy use was positively predicted by mastery and performance climates (Ommundsen, 2006; Theodosiou et al., 2008) and mastery goals (Theodosiou et al., 2008). Regulation of effort was positively predicted by performance-approach goals, mastery climates, and performance climates, and negatively predicted by performance-avoidance goals (Ommundsen, 2006). Help-seeking was positively predicted by mastery and performance climates (Ommundsen, 2006). Self-handicapping was positively predicted by performance-avoidance goals and performance climates (Ommundsen, 2004, 2006), and negatively predicted by mastery goals (Ommundsen, 2004) and performance-approach goals (Ommundsen, 2004, 2006). Mastery goals mediated the self-handicapping reducing effect (Ommundsen, 2001).

Fourteen studies examined the predictive utility of AGT constructs on heterogeneous cognitive responses. Students' positive attitude toward daily physical activity at school was positively related to mastery goals, mastery and performance climates (Halvari et al., 2011). Students' positive attitude toward PE class was positively predicted by profiles high in mastery climate and moderate in performance climate, and students' negative attitude toward PE class was negatively predicted by climate profiles high in performance climate and low in mastery climate (Treasure, 1997). Students' intentions to be physically active were positively related to mastery goals (Sproule et al., 2007), and students' intentions to run were positively predicted by performance goals in boys (Xiang et al., 2004b) and high scores in mastery goals (Xiang et al., 2007). Responsibility and relationship goals were positively related to mastery climates (Fernández-Rio et al., 2014; Baena-Extremera et al., 2015b) and performance climates (Baena-Extremera et al., 2015b). Mastery climate was positively related to perceptions of teaching strategies promoting students' internal locus of causality, and performance climate to promoting students' external locus of causality for the regulation of disciplined behavior (Spray, 2002). Perceptions of effort and ability as causes of success positively matched with profiles high in mastery climate and moderate in performance climate, and perceptions of ability as cause of success were positively related to profiles high in performance climate and low in mastery climate (Treasure, 1997). Students' physical self-worth and global self-esteem were positively related to profiles high
in mastery and performance goals (Kavussanu, 2007). Perceived usefulness of PE classes was positively related to performance climates in girls (Baena-Extremera et al., 2013). Positive aspects of students’ ethical cultural salience were positively related to mastery climates and mastery goals, while negative aspects were positively related to performance climates and performance goals (Kouli and Papaioannou, 2009). Situational interest was positively related to mastery goals and performance-avoidance goals (Shen et al., 2007). Finally, knowledge gain in softball and self-reported metacognitive knowledge were positively related to mastery climates (Shen et al., 2007; Theodosiou et al., 2008).

Three studies addressed contextual outcome variables of PE students’ teacher initiated perceived motivational climate:

- Exploring the association between school- and classroom goal structures, one study (Barkoukis et al., 2012) suggested that school goal structures predict the respective structures within the PE classroom. Here, students’ perceptions of a mastery school goal structure turned out to be both positive and negative predictors of mastery and performance classroom structures, respectively. Performance school goal structures proved to be positive predictors of performance classroom structures.

- Two studies (Papaioannou, 1998; Spray, 2002) explored the relationships among students’ perceptions of different PE motivational climates and perceived teacher strategies to sustain class control and discipline. In both studies, perceived mastery climate was positively linked to perceived teaching strategies promoting more student-determined reasons for exercising discipline. On the other hand, performance climate positively matched to students’ perceptions of PE teachers accentuating introjected reasons of controlling the classroom as well as the perception of teachers not being concerned with class discipline.

### Intervention Studies

Table 2 summarizes the descriptive characteristics of the included intervention studies. Publication dates ranged from 1996 to 2017. All studies but two (Erturan-Ilker and Demirhan, 2013; Erturan-Ilker, 2014) were completed in Europe or the USA. In total, 4,595 participants were included in the 21 studies, of which at least 2,084 were female (about 51.9%; one study made no specification about gender). The average sample size was 128, comprising a sample range between 35 and 782 students. Mostly the participants were between 11 and 16 years old, with only two studies focusing on younger participants (Todorovich and Curtner-Smith, 2003; Logan et al., 2014). All studies but two (Erturan-Ilker and Demirhan, 2013; Erturan-Ilker, 2014) were based in the dichotomous framework of AGT. Perceived achievement goals were primarily gathered by the TEOSQ (n = 7; 53.9%) and the POSQ (n = 3; 23.1%). To gauge the perceived motivational climate, the LAPOPECQ (n = 5; 31.3%) and the PMCS (n = 4; 25%) were mainly utilized.

The majority of the studies examined the effects of interventions based on the TARGET areas to create a mastery and/or performance-involving motivational climate during PE lessons (n = 16). Two studies examined the climate effects of positive and negative feedback (Viciana et al., 2007; Erturan-Ilker, 2014). Three studies (Christodoulidis et al., 2001; Standage et al., 2007; Erturan-Ilker and Demirhan, 2013) reported the use of other teaching interventions aimed at profitably manipulating PE motivational climates. The interventions were mostly conducted by regular PE teachers (n = 15) with the exception of five studies where the interventions were conducted by the researchers and one study by graduate students. In 13 studies, intervention groups were compared to control groups that received no intervention and were taught by teachers using their typical teaching style (Christodoulidis et al., 2001; Morgan and Carpenter, 2002; Todorovich and Curtner-Smith, 2002, 2003; Weigand and Burton, 2002; Digelidis et al., 2003; Jaakkola and Liukkonen, 2006; Barkoukis et al., 2008, 2010; Almolda-Tomás et al., 2014; Sevil et al., 2015; Abós et al., 2016; Papaioannou et al., 2017). Intervention and control groups were mostly examined by different teachers or researchers (Christodoulidis et al., 2001; Morgan and Carpenter, 2002; Todorovich and Curtner-Smith, 2002, 2003; Digelidis et al., 2003; Jaakkola and Liukkonen, 2006; Barkoukis et al., 2008, 2010; Almolda-Tomás et al., 2014; Sevil et al., 2015; Abós et al., 2016; García-González et al., 2017). In many cases, control groups were recruited from different schools than the school in which the intervention took place. Only in one case did the same teachers or researchers teach both intervention and control groups (Weigand and Burton, 2002). The contents of the experimental and control group lessons were either the same (Morgan and Carpenter, 2002; Weigand and Burton, 2002; Barkoukis et al., 2008, 2010; Almolda-Tomás et al., 2014; Sevil et al., 2015; Abós et al., 2016; Garcia-González et al., 2017) or different (Christodoulidis et al., 2001; Todorovich and Curtner-Smith, 2002, 2003; Digelidis et al., 2003; Jaakkola and Liukkonen, 2006). Six studies compared two or more intervention groups and had no control group (Solmon, 1996; Standage et al., 2007; Viciana et al., 2007; Erturan-Ilker and Demirhan, 2013; Erturan-Ilker, 2014; Logan et al., 2014). Two studies used within-subject designs: Bortoli et al. (2017) compared two experimental groups that received consecutively both a mastery and a performance climate intervention but in reverse sequence, and Papaioannou and Kouli (1999) compared a lesson with mastery-involving tasks with a lesson with performance-involving tasks. The duration of the interventions largely varied across studies. The shortest intervention period covered one running task (Standage et al., 2007), and two PE lesson interventions in volleyball or juggling (Solmon, 1996; Papaioannou and Kouli, 1999). The longest interventions were applied in one academic year (Christodoulidis et al., 2001; Digelidis et al., 2003). Mostly, the PE lessons took place twice a week. Only four studies reported other frequencies: once a week (Almolda-Tomás et al., 2014), three times a week (Digelidis et al., 2003), and daily PE lessons (Todorovich and Curtner-Smith, 2002, 2003). The majority of studies used a pre-post design and compared the dependent variables before and after the intervention. Two
studies each added an intermediate test (Papaioannou and Kouli, 1999; Bortoli et al., 2017) or reported on follow-up results (Christodoulidis et al., 2001; Digelidis et al., 2003).

Supplementary Table 6 provides a study-by-study overview of the analytical variables and main outcomes. The substantial results are summarized below by means of three categories (see Figure 4). The categories again were qualitatively developed and based on inductively grouped key outcome topics regarding the examined intervention effects.

1. A first topic considered AGT outcomes as dependent variables and examined intervention effects on PE students' perceived achievement goals and perceived motivational climates (n = 16):

- TARGET-based interventions designed to implement mastery-involving climates led to (a) higher levels of perceived mastery goals (Solmon, 1996; Morgan and Carpenter, 2002; Todorovich and Curtner-Smith, 2002, 2003; Weigand and Burton, 2002; Digelidis et al., 2003; Jaakkola and Liukkonen, 2006; García-González et al., 2017), partially without effect of gender (Solmon, 1996; Todorovich and Curtner-Smith, 2002, 2003) and partially with girls reporting more mastery goals than boys (Jaakkola and Liukkonen, 2006), (b) decreases in perceived performance goals (Digelidis et al., 2003; Jaakkola and Liukkonen, 2006), (c) higher levels of perceived mastery climates (Papaioannou and Kouli, 1999; Morgan and Carpenter, 2002; Digelidis et al., 2003; Sevil et al., 2015; Abós et al., 2016; García-González et al., 2017), with one study outlining positive effects for both gender and slight advantages for boys (Sevil et al., 2015), and (d) lower values on perceived performance climate (Papaioannou and Kouli, 1999; Digelidis et al., 2003; Sevil et al., 2015; Abós et al., 2016; García-González et al., 2017).

- The studies implementing alternative mastery-climate manipulations confirmed these results regarding perceived mastery goals (Erturan-Ilker and Demirhan, 2013) and both perceived motivational climates (Christodoulidis et al., 2001). To this end, it is remarkable that in both studies which conducted follow-up measurements, nearly all previous in-between group differences dispersed (Christodoulidis et al., 2001; Digelidis et al., 2003). TARGET-based (Solmon, 1996; Todorovich and Curtner-Smith, 2002, 2003) and other (Erturan-Ilker and Demirhan, 2013) performance-oriented climate treatments both resulted in higher levels of perceived performance goals, partially without effect of gender (Todorovich and Curtner-Smith, 2002, 2003) and partially with girls reporting more mastery goals than boys (Solmon, 1996).

- Interventions using positive feedback strategies led to higher values on perceived mastery climates and mastery goals (Viciana et al., 2007; Erturan-Ilker, 2014), and decreasing scores on perceived performance-avoidance climate and goals (Erturan-Ilker, 2014). Negative feedback strategies eventuated in higher scores on perceived performance climates and goals (Viciana et al., 2007), performance-avoidance climate and goals, and decreasing values on perceived mastery climate (Erturan-Ilker, 2014).

2. A second focus was the examination of TARGET-based treatment effects on SDT outcomes (n = 7). Interventions structuring a mastery-involving motivational climate in PE lessons enhanced BPN such as perceived competence (Weigand and Burton, 2002; Barkoukis et al., 2008; Almolda-Tomás et al., 2014; Sevil et al., 2015; Abós et al., 2016), perceived autonomy (Almolda-Tomás et al., 2014; Sevil et al., 2015; Abós et al., 2016), and perceived relatedness (Abós et al., 2016), and showed valuable effects on different elements of self-determined motivation. Students within the mastery groups reported higher levels of self-determination with boys perceiving less self-determined motivation than girls (Jaakkola and Liukkonen, 2006), intrinsic motivation (Sevil et al., 2015) and identified regulation (Almolda-Tomás et al., 2014; Sevil et al., 2015), and lower levels of amotivation (Jaakkola and Liukkonen, 2006; Almolda-Tomás et al., 2014). Using a treatment combining performance and mastery climates, the intervention of Bortoli et al. (2017) effectively changed self-determined motivation in the performance-mastery group, with the associated students reporting lower scores on intrinsic motivation and higher scores on amotivation after the initial performance phase. However, the levels of self-determined motivation within this group increased during the mastery climate phase but remained below the level before the performance phase.

3. A third core theme emphasized intervention effects on further psychological outcomes (n = 16):

- Regarding affective variables, TARGET-based interventions supportive of a PE mastery climate led to higher levels of satisfaction and enjoyment (Morgan and Carpenter, 2002; Weigand and Burton, 2002; Barkoukis et al., 2008; Almolda-Tomás et al., 2014; Sevil et al., 2015), lessened boredom (Weigand and Burton, 2002), worry (Barkoukis et al., 2008), and somatic anxiety (Papaioannou and Kouli, 1999). Positive feedback focused on individual ability and effort affirmed these outcomes for enjoyment (Viciana et al., 2007).

- For cognitive consequences, the TARGET-based interventions affected positive attitude and predisposition toward PE, sports, healthy eating (Morgan and Carpenter, 2002; Digelidis et al., 2003; Abós et al., 2016), and preferences in difficult task choices (Morgan and Carpenter, 2002). Studies using alternative mastery-climate manipulations supported these results (Digelidis et al., 2003; Erturan-Ilker and Demirhan, 2013). Conversely, students in TARGET-based performance climates were more likely to attribute success to ability with boys rating ability higher than girls (Solmon, 1996). Performance climate instructions served as positive predictors of students’ situational self-handicapping (Standage et al., 2007). Negative feedback focused on individual ability and effort gave rise to higher preferences for easy task choices (Viciana et al., 2007).
<table>
<thead>
<tr>
<th>Source</th>
<th>N (girls)</th>
<th>Mean age + SD</th>
<th>Nation</th>
<th>School</th>
<th>Study design</th>
<th>Intervention program</th>
<th>AGT frame</th>
<th>AGT scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abós et al., 2016</td>
<td>IG 15 (9) CG 20 (10)</td>
<td>IG 15.40 ± 0.51 CG 15.35 ± 0.49</td>
<td>Spain</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); pre-post-measurement of predisposition toward PE/post measurement of BPN/mot. climate</td>
<td>6 weeks (12 lessons, two per week, 55 min per lesson); based on TARGET areas; acrosport</td>
<td>Dicho</td>
<td>PMCS</td>
</tr>
<tr>
<td>Almolda-Tomás et al., 2014</td>
<td>113 (53)</td>
<td>IG 14.83 ± 0.72 CG 14.84 ± 0.68</td>
<td>Spain</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); post-measurement</td>
<td>4 weeks (four double sessions, one per week, 100 min per double session); based on TARGET areas</td>
<td>Dicho</td>
<td>PMCS</td>
</tr>
<tr>
<td>Barkoukis et al., 2010</td>
<td>IG 131 (74) CG 186 (96)</td>
<td>IG 13.9 ± 0.61 CG 13.9 ± 0.81</td>
<td>Greece</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); pre-post measurement</td>
<td>10 consecutive sessions; based on TARGET areas; athletics</td>
<td>Dicho</td>
<td></td>
</tr>
<tr>
<td>Barkoukis et al., 2008</td>
<td>IG 183 (86) CG 181 (94)</td>
<td>IG 13.9 ± 0.81 CG 13.8 ± 0.69</td>
<td>Greece</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); post-measurement</td>
<td>7 month (regular PE lessons); based on TARGET areas; games, athletics, gymnastics and traditional dances</td>
<td>Dicho</td>
<td>LAPOPECO, TEOQS</td>
</tr>
<tr>
<td>Bortoli et al., 2017</td>
<td>65 (65)</td>
<td>14–15</td>
<td>Italy</td>
<td>Sec</td>
<td>NRES; two groups (MC/PC, PC/MC); pre-intermediate-post measurement</td>
<td>Two intervention phases (one group with MC in phase 1/PC in phase 2, one group in reversed order; 4 month per phase; eight lessons per phase); based on TARGET areas; basic gymnastic tasks</td>
<td>Dicho</td>
<td>TIMCEPQ</td>
</tr>
<tr>
<td>Christodoulidis et al., 2001</td>
<td>IG 105 (58) CG 29 (340)</td>
<td>15–16</td>
<td>Greece</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); pre-post measurement, follow-up (10 month after end of intervention)</td>
<td>One academic school year (regular PE lessons); self-designed teaching intervention (resulting in 25 analytic daily lesson plans)</td>
<td>Dicho</td>
<td>LAPOPECO, TEOQS</td>
</tr>
<tr>
<td>Digelidis et al., 2003</td>
<td>IG 262 (132) CG 20 (276)</td>
<td>IG 11.88 ± 0.60 CG 12.14 ± 0.77</td>
<td>Greece</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); pre-post measurement, follow-up (10 month after end of intervention)</td>
<td>One academic school year (88 lessons, three per week, 45 min per lesson); based on TARGET areas</td>
<td>Dicho</td>
<td>LAPOPECO, TEOQS</td>
</tr>
<tr>
<td>Erturan-Iker and Demirhan, 2013</td>
<td>81 (40)</td>
<td>14.38 ± 0.49</td>
<td>Turkey</td>
<td>Sec</td>
<td>NRES; three groups (MC, PAp,PAv); pre-post measurement</td>
<td>12 weeks; based on MCOCL; fundamental practices, athletics, volleyball, basketball</td>
<td>Tricho</td>
<td>MCOCL, TAGS</td>
</tr>
<tr>
<td>Erturan-Iker, 2014</td>
<td>PF 27 (15) NF 20 (12)</td>
<td>PF 15.62 NF 15.74</td>
<td>Turkey</td>
<td>Sec</td>
<td>NRES; two groups (PFNP); pre-post measurement</td>
<td>6 weeks (90 min per week); volleyball</td>
<td>Tricho</td>
<td>TMCS, TAGS</td>
</tr>
<tr>
<td>García-González et al., 2017</td>
<td>IG 263 CG 317</td>
<td>12.46 ± 1.15</td>
<td>Spain</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); post-measurement (after last session of each unit)</td>
<td>Four teaching units (each lasting 10 sessions; 2 per week, 50 min per session); based on TARGET areas; long-distance running, rope skipping, sport orienteering, volleyball</td>
<td>Dicho</td>
<td>TMCS</td>
</tr>
<tr>
<td>Jaakkola and Luukkanen, 2006</td>
<td>IG 178 (87) CG 59 (123)</td>
<td>15–16</td>
<td>Finland</td>
<td>Sec</td>
<td>NRES; two groups (IG,CG); pre-post measurement</td>
<td>One academic school year (regular PE lessons, ~ 30 lessons, 90 min per lesson); based on TARGET areas</td>
<td>Dicho</td>
<td>POSQ</td>
</tr>
<tr>
<td>Logan et al., 2014</td>
<td>MC 23 (13) PC 25 (12)</td>
<td>7.8</td>
<td>US</td>
<td>Elem</td>
<td>RCT; two groups (MC,PC); pre-post measurement</td>
<td>5 weeks (daily lessons, 25 min per lesson); based on TARGET areas; basic motor tasks, skills and games</td>
<td>Dicho</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Source</th>
<th>N (girls)</th>
<th>Mean age ± SD</th>
<th>Nation</th>
<th>School</th>
<th>Study design</th>
<th>Intervention program</th>
<th>AGT frame</th>
<th>AGT scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morgan and Carpenter, 2002</td>
<td>153 (54)</td>
<td>13.6 ± 0.64</td>
<td>UK Sec</td>
<td>NRES; two groups (IG, CG); pre-post measurement (except motivational climate: only post)</td>
<td>7 weeks; based on TARGET areas; athletics</td>
<td>Dicho</td>
<td>LAPOPECQ, TEOSQ</td>
<td></td>
</tr>
<tr>
<td>Papaioannou and Kouli, 1999</td>
<td>239 (108)</td>
<td>13 ± 0.50</td>
<td>Greece Sec</td>
<td>NRES; within-subject comparison; pre-intermediate-post measurement</td>
<td>Two lessons at 14-day intervals (45 min per lesson; one with mastery, one with performance oriented drills); based on TARGET areas; volleyball</td>
<td>Dicho</td>
<td>LAPOPECQ, TEOSQ</td>
<td></td>
</tr>
<tr>
<td>Sevil et al., 2015</td>
<td>IG 109 (58)</td>
<td>12.37 ± 0.64</td>
<td>Spain Sec</td>
<td>NRES; 2 groups (IG, CG); post measurement</td>
<td>3 month (10 lessons, 2 per week, 50 min per lesson), based on TARGET areas; rope skipping</td>
<td>Dicho</td>
<td>PMCS</td>
<td></td>
</tr>
<tr>
<td>Solmon, 1996</td>
<td>109 (53)</td>
<td>Not reported US Sec</td>
<td>RCT; 2 groups (MC, PC); post measurement</td>
<td>Two lessons (30 min per lesson); based on TARGET areas; juggling</td>
<td>Dicho</td>
<td>PMCSQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standage et al., 2007</td>
<td>MC 34 (20)</td>
<td>11.98 ± 0.31</td>
<td>UK Sec</td>
<td>RCT; 2 groups (MC, PC); pre measurement of DV, post measurement of mot. climate</td>
<td>One experimental task; based on MC/PC instructions; 20 m shuttle run test</td>
<td>Dicho</td>
<td>POSQ</td>
<td></td>
</tr>
<tr>
<td>Todorovich and Curtner-Smith, 2002</td>
<td>MC 23 (10)</td>
<td>11 ± 3.3 mo.</td>
<td>US Elem</td>
<td>RCT; 3 groups (MC, PC, CG); pre-post measurement</td>
<td>2 weeks (daily lessons, 30 min per lesson); based on TARGET areas; field hockey (MC/PC), softball (CG)</td>
<td>Dicho</td>
<td>PECAI, TEOSQ</td>
<td></td>
</tr>
<tr>
<td>Todorovich and Curtner-Smith, 2003</td>
<td>MC 28 (13)</td>
<td>Not reported</td>
<td>US Elem</td>
<td>RCT; 3 groups (MC, PC, CG); pre-post measurement</td>
<td>2 weeks (daily lessons, 30 min per lesson); based on TARGET areas; field hockey (MC/PC), softball (CG)</td>
<td>Dicho</td>
<td>PECAI, TEOSQ</td>
<td></td>
</tr>
<tr>
<td>Viciana et al., 2007</td>
<td>95 (44)</td>
<td>14–16</td>
<td>Spain Sec</td>
<td>NRES; 3 groups (PF, NF, BF); pre-post measurement</td>
<td>7 weeks (14 lessons, Two per week); ~6 feedback interventions per student per lesson; acrosport, pyramid constructions</td>
<td>Dicho</td>
<td>LAPOPECQ, TICO, POSQ</td>
<td></td>
</tr>
<tr>
<td>Weigand and Burton, 2002</td>
<td>40 (8)</td>
<td>15.9 ± 0.51</td>
<td>UK Sec</td>
<td>RCT; 2 groups (MC, CG); pre-post measurement</td>
<td>5 weeks (10 lessons, 60 min per lesson); based on TARGET areas; soccer techniques</td>
<td>Dicho</td>
<td>TEOSQ</td>
<td></td>
</tr>
</tbody>
</table>

BF, group that received both positive and negative feedback; CG, control group; elem, elementary education; IG, intervention group; LAPOPECQ, Learning and Performance Orientations in PE Classes Questionnaire; MC, mastery climate group; MCOCL, Motivational Climate Observer Control List; NRES, non-randomized experimental studies; PAp, performance-approach group; PAv, performance-avoidance group; NF, negative feedback group; PC, performance climate group; PECAI, PE Climate Assessment Instrument; PF, positive feedback group; PMCS, Perceived Motivational Climate Scale; PMCSQ, Perceived Motivational Climate in Sport Questionnaire; POSQ, Perception of Success Questionnaire; RCT, randomized controlled trial; sec, secondary education; SOFIT, System for Observing Fitness Instruction Time; TAGS, Trichotomous Achievement Goal Scale; TEOSQ, Task and Ego Orientation in Sport Questionnaire; TICO, Teacher-initiated Competitive Orientation Scale; TIMCPEQ, Teacher-Initiated Motivational Climate in PE Questionnaire; TMCS, Trichotomous Motivational Climate Scale.
FIGURE 4 | Structuring categories for included intervention studies. This figure depicts the different categories that were used in organizing the teaching intervention studies based on outcomes.

- Considering behavioral consequences as dependent variables, mastery-involving TARGET-climates occasioned more time into health-related moderate-to-vigorous physical activity (Logan et al., 2014), and less time in management tasks such as class business, behavior management, and time spent in transition. Additionally, mastery-involved students completed higher numbers of difficult practice trials per minute (Solmon, 1996), and showed better technical executions in athletics (Barkoukis et al., 2010). Using a self-designed teaching intervention, Christodoulidis et al. (2001) reported increased exercise timespans.

- For motivational responses, TARGET-based and alternative mastery-involving treatments showed higher values in sports-related self-efficacy (Barkoukis et al., 2010) and motivational strategies (Erturan-Iker and Demirhan, 2013), respectively. Implementing psychobiosocial (PBS) variables, the study of Bortoli et al. (2017) ran counter to the outcome categories listed above. As for SDT constructs, the intervention effectively changed PBS states in the performance-mastery group, with the associated students reporting lower scores on pleasant/functional BPS states after the initial performance phase, and lower scores on pleasant/functional BPS states after the second mastery phase.

DISCUSSION

The guiding aim of this paper was to substantiate established justification strategies of PE by theory and empirical evidence of AGT. Following theoretical considerations confirming the pedagogical value of AGT, a systematic review of the international literature was conducted. The final sample included 91 trials. The majority of the investigations implemented cross-sectional study designs and thus stabilized a typical methodological feature of achievement goal research. At the same time, various scientific demands for enhanced AGT-research on long-term effects and experimental approaches (e.g., Biddle et al., 2003; Valentini and Rudisill, 2006; Braithwaite et al., 2011) seem increasingly to be heard and, regarding the ratio of included study designs in the last 10 years, suggest at least converging establishment of cross-sectional studies, longitudinal studies, and intervention studies. Relative to the structure of the results section, the main findings of the included observational and intervention studies are briefly summarized below, classified by content and methodologically reflected.

Observational Studies

Observational study designs represent the standard of research in educational psychology. Since the independent variable is thereby not manipulated by the researcher(s), the obtained study results occupy a comparatively low position in the hierarchy of scientific evidence (Joanna Briggs Institute [JBI], 2018). Nevertheless, the current review involved 70 observational studies exploring AGT in PE. Most studies applied a cross-sectional design, only a few studies applied a longitudinal design. In total, the studies provided important findings concerning group differences, interrelations, predictors, and outcomes of various achievement goals and motivational climates in PE lessons (see Figure 3):

1. Group differences within AGT constructs primarily focus on gender and age. Although the reported results differed between the studies, a predominant trend was observed with females and younger students being more mastery-oriented in terms of achievement goals and perceived motivational climate while males and older students were more performance-oriented. Despite this trend, it is important to note that some studies reported no gender or age differences, and that the variance between studies suggests that further research should be conducted to derive conclusive information on gender, age, and especially on less researched country and school-differences of students. Moreover, further
research should consider not only achievement goals but also motivational climate.

2. Concerning interrelations, numerous findings highlighted mutual alignment of personal and contextual aspects of AGT. Mastery goals were often positively related to mastery climate, while performance goals were positively related to performance climate. Furthermore, mastery goals were positively related to adaptive psychological variables, and negatively related to maladaptive variables. Findings surrounding performance goals were less conclusive, but motivational climate appears to be positively related to maladaptive variables, and negatively to adaptive variables.

3. Only a few studies placed a focus on personal and contextual predictor variables of AGT constructs. Incremental beliefs of ability, internal explanations for events, high levels of perceived competence, self-determined types of motivational regulation, and mastery climates emphasized by teachers were the main positive predictors for mastery and performance-approach goals. Entity beliefs of ability, external explanations for events, low levels of perceived competence, controlled types of motivational regulation, and performance climates emphasized by peers were reported to be positive prognostic conditions for performance and performance-avoidance goals mainly.

4. Most observational studies positioned achievement goals and motivational climates as predictors of different outcomes. The results mostly confirm the previous statements considering the interrelation studies. For a variety of affective, cognitive, behavioral, and motivational variables it applies that mastery goals, mastery-approach goals, and mastery climates are predominantly related to more positive consequences. In contrast, performance goals, performance-avoidance goals, and performance climate are widely associated with negative responses. Some findings were unique in that performance-approach goals were associated with rather adaptive outcome patterns. In addition, analyses using goal profiles showed that moderate to high levels of mastery goals can “buffer” or even positively affect levels of performance goals.

Content

For systematic reviews it is not meaningful to discuss the summarized results contentwise in comparison to primary studies. Therefore, the contentwise discussion is presented in relation to the state of research in similar research areas focusing on physical activity and school subjects:

- Systematic reviews in other research areas of physical activity widely confirm our results. Mastery goals and mastery climates in sports, exercise, youth sports, and competitive sports are fundamentally related to more adaptive consequences. Performance goals and performance climates are associated with rather maladaptive responses (Ntoumanis and Biddle, 1999; Harwood and Biddle, 2002; Biddle et al., 2003; Harwood et al., 2008; Lochbaum et al., 2016). Furthermore, studies focusing on achievement goal profiles substantiate the conceptual orthogonality of achievement goals (Nicholls, 1989) and the pioneering importance of mastery goals. To this end, coupled with high or moderate mastery goals, high scores in performance goals do not necessarily lead to negative outcomes (Duda and Whitehead, 1998; Smith et al., 2006; Roberts, 2012). In performance-orientated sport contexts and their agonal orientation toward victorious competitions, partially positive effects were found in performance-approach goals (e.g., Elliot and Conroy, 2005).

- The state of research for other school subjects depicts a similar picture, although because of nearly no existing systematic reviews, primary studies have to be considered for comparison. Mastery goals and mastery climates are often positively associated with subject-related academic achievement and psychological outcomes such as motivation, interest, and self-efficacy. The results of performance goals very often point in the opposite direction or do not show a predicting function (Nicholls et al., 1990; Zusho et al., 2005; Matos et al., 2007; Lau and Nie, 2008; Linnenbrink-Garcia et al., 2008; Keys et al., 2012; Pantziara and Philippou, 2014). In some studies, students with a high level of performance-approach goals in their goal profile show higher academic achievement than students with other goal profiles (Valle et al., 2003; Bong, 2009). Additionally, empirical evidence shows that students pursue similar achievement goals in different school subjects and that achievement goal motivation shows a similar longitudinal development in different school subjects (Duda and Nicholls, 1992; Bong, 2001; Hornstra et al., 2016, 2017). However, these existing comparative studies did not account for PE.

In sum, taking both comparison areas into account, a certain level of generality or trans-contextuality can be seen for both achievement goals and motivational climates. With respect to motivational theory, a potential outlying or illegitimizing status of PE within the context of physical activity and other school subjects is therefore at most marginal.

Methodology

Several tendencies can be observed for the observational studies with respect to methods and methodology. Most studies applied a cross-sectional design while only a few studies applied a longitudinal design. Therefore, causal conclusions are very limited. Appropriate and advanced statistical approaches have been applied in many of the included studies (e.g., structural equation models or hierarchical multiple regression techniques). Therefore, it seems that researchers in the field of educational psychology are highly aware of the necessity to carefully account for complex school structures with adequate statistical techniques. Most studies analyzed PE classes in secondary education classes. Therefore, information regarding younger students in primary schools is rather limited and further research is warranted. The lowest number of participants is $n = 105$ and many studies were able to include several 100 and up to nearly 3,000 students. The applied study designs therefore, seem to be most often appropriate to draw reliable results based on the sample size.
**Intervention Studies**

Using experimental study designs and controlled conditions, intervention studies are prospectively and analytically configured, and promise causal significance and high levels of scientific evidence. However, compared to the amount of observational study designs, the amount of included intervention studies was relatively low. The review comprised 21 intervention studies that manipulated various contextual determinants of students’ mastery and performance involvement in PE (see Figure 4). The implemented manipulations of the PE motivational climate were based on the TARGET strategies, self-designed teaching programs and different feedback strategies. The findings demonstrated a distinct trend across the interventions:

- The interventions were successfully implemented and fulfilled their purpose to practically incorporate mastery or performance climates within PE lessons in all studies.
- Mastery and performance-oriented climate manipulations clearly led to perceived mastery and performance-oriented climates, respectively. PE environments aligned to self-referenced criteria of success and failure, and value demonstrations of mastery and learning were consistently associated with both pedagogically adaptive effects and curricularly intended consequences. These motivational and psychological advantages were thereby shown over regular PE lessons without special focus on achievement motivational aspects as well as in PE environments characterized by other-referenced criteria of success and failure, and normative ability. In interventions emphasizing a performance climate, additionally a broad range of absent or maladaptive treatment effects became apparent.

**Content**

Contentwise, these findings confirm the trans-contextual persistence of AGT which has already been discussed for the observational studies at a higher level of evidence:

- For different fields of physical activity, previous syntheses regarding specifically TARGET-based interventions in PE (Braithwaite et al., 2011) and motor skill interventions in children with and without developmental motor delays (Valentini and Rudisill, 2006; Palmer et al., 2017; Ribeiro Bandeira et al., 2017) draw a comparable picture. For other fields of physical activity, according to our knowledge, there are no overviews. However, the use of individual studies shows similarly consistent findings for sports in general, youth sports, competitive sports and preventive sports programs (e.g., Theboom et al., 1995; Smith et al., 2007; Conde et al., 2009; Maro et al., 2009; Cecchini et al., 2014; Hassan and Morgan, 2015; McLaren et al., 2015). For a final assessment and due to a swiftly increasing number of AGT intervention studies in all areas of physical activity, further systematic reviews and meta-analyses are both necessary and desirable.
- For other school subjects, very few intervention studies could be ascertained. Ames (1992a) pointed to more positive attitudes toward mathematics and school per se in an one-year intervention with elementary teachers and students. In a study examining effects of a quasi-experimental classroom goal condition, Linnenbrink (2005) revealed differences on help-seeking and achievement with the combined mastery/performance-approach condition showing the most beneficial situational pattern. The scientific application of AGT in academic contexts thus seems to specifically focus on PE.
- In addition, the findings regarding gender effects (Solmon, 1996; Papaioannou and Kouli, 1999; Todorovich and Curtner-Smith, 2002, 2003; Jaakkola and Liukkonen, 2006; Sevil et al., 2015) contrast studies with and without group differences. Further studies are needed here to evaluate the personal and situational backgrounds of these inconsistencies.

**Methodology**

Regarding methods and methodology, four main tendencies can be observed for the intervention studies. First, the uneven distribution of experimental and quasi-experimental study designs is remarkable. All included research consists of field investigations situated in real-world PE contexts. Apart from some exemptions (Solmon, 1996; Todorovich and Curtner-Smith, 2002, 2003; Weigand and Burton, 2002; Standage et al., 2007; Logan et al., 2014), most studies recruited their experimental groups from existing school structures and class groups. The allocation to study groups therefore is at most limited random and harbors the risk of selection bias. However, the methodological silver bullet of intervention research is the traditional experiment, implemented as RCTs. This also applies for intervention research within the field of education (Leutner, 2010). Against the much-discussed claim that it is impossible to implement RCTs in education (Slavin, 2002; Andrews, 2005), a recent review of 1,017 RCTs in education research systematically confirmed the educational feasibility of RCTs, in particular for school settings and school research objects (Connolly et al., 2018). Second, the included studies are almost exclusively limited to post intervention measurements in the immediate proximity to the ends of the interventions. Thus, almost no statements are possible about how long the effects last and how stable they are. Third, the teaching manipulations are constructed by a variety of components. The majority of the studies implemented an intervention using the TARGET strategies. The self-designed interventions combined alternative didactic and methodical elements. Due to the large number of modules and starting points, however, it remains fairly unclear which components of the interventions led to the identified effects. Thus, firm conclusions on the positive effects of motivational climate interventions and practical implications can only be derived as complex “wholes”. Evidence-based individual statements and practical starting points can therefore only be derived to a limited extent at the moment (for similar criticisms see Braithwaite et al., 2011; Cecchini et al., 2014). Fourth, for the selection and recording of outcomes, a similar ambivalence is shown. Although the studies uniformly highlight the importance of AGT for PE and physical activity beyond PE, in only a few
cases variables relating to the practices, content, and goals of PE were included as dependent variables. To this end, with the exception of the study by Logan et al. (2014), physical activity is captured merely subjectively. Statements considering objective effects are therefore limited. Objective measurement of physical activity using ambulatory assessment could be a promising methodological supplement. In sum, it is therefore not yet possible to derive sophisticated recommendations regarding duration, frequency, and PE contents based on the results. Future studies should start here in order to reinforce the school subject-justifying usefulness of AGT in more detail and to enable more concrete practical implications.

**Strengths and Limitations**

The major strength of this paper is the targeted combination of theoretical reasoning and empirical data. In order to achieve the objectives, in a first step we presented the pedagogical PE value of AGT theoretically, and in a second step we confirmed this argumentative foundation by results of empirical studies exploring motivation in PE from an AGT viewpoint. The empirical studies are therefore synthesized in a systematic review according to the PRISMA statement. We applied a diversified search in recognized online databases. The search terms were guided by theoretical reasoning. The choices of databases were grouped and cover a range of relevant scientific disciplines. Two reviewers independently conducted the study selection and assessed the risk of bias of the included studies. Approved appraisal tools were used for the quality assessment. Thus, the results are widely replicable and methodically indicate significance due to a high position of systematic reviews within the scientific hierarchy of evidence (Joanna Briggs Institute [JBI], 2014). However, some limitations warrant attention. First, concerning AGT, a differentiated range of concepts and conceptual differences emerged since the foundational theoretical works (Dweck, 1984; Nicholls, 1984). In particular, for the performance approach and performance-avoidance goals, a current understanding has developed that additionally distinguishes between normative, appearance, and evaluative aspects (Hulleman et al., 2010; Daumiller et al., 2019).

Although our review acknowledges these conceptualizations in the theoretical section, within the presentation of included studies it defines performance goals on a more overall and undifferentiated level to counteract an even smaller stepping of the already complex representation of results. Second, only peer-reviewed journal articles published in English language and listed in the screened databases were included. Studies in additional languages, gray literature or reports, and articles outside the scope of the selected databases went unheeded. The existence of relevant studies in other languages or studies published elsewhere therefore cannot be disclosed. Third, any systematic review is not able to discriminate between low reporting and low methodological quality of studies and hence low values of methodological quality may either reflect weak reporting or weak study designs. Fourth, due to relatively low methodological quality ratings, comparatively many longitudinal studies were excluded during the study selection process. A bias in the number of longitudinal studies therefore has to be considered.

**CONCLUSION**

School PE is an inherent but disputed component of international school systems. Unlike most school subjects, PE consistently has to vindicate its schooling authority and has therefore designed differentiated justification strategies, to justify the pedagogical value of PE by a canon of psychomotor, moral, affective, cognitive, and behavioral aspects. The present study takes up these rationales and asks for the empirically saturated contribution of AGT to these conceptually desired PE outcomes. To conclude, the results provide comprehensive psychological backup to plan and shape PE lessons matching the intended curricular ambitions. PE students' perceived achievement goals and perceived PE motivational climates are in multiple relations to desired or undesired consequences about, in, of, and through the physical, and often prove to be explanatory variables for the occurrence of these consequences. Perceived mastery goals and mastery climates are of particular pedagogical significance and positively touch on multiple curricular PE aims such as motor skill development, sports participation beyond the boundaries of the PE classroom, psychological well-being, and prosocial dispositions or aspects of healthy living. Profiles combining high mastery goals and high performance or performance-approach goals partly also show an expedient character. However, any productive integration of these results into everyday school PE practice seems partially questionable due to an at least partially prevalent performance-pedagogical attitude of PE teachers and a popular PE-orientation toward sports, competition, and social comparison (e.g., Wolters, 2012; Coulter and Chróinín, 2013; Svendsen and Svendsen, 2016; Schierz and Serwe-Pandrick, 2018; Stirrup, 2018). For a practical meaning of these justifying results, a change of this basic PE-teacher habitus therefore seems to be both indispensable and promising.

**AUTHOR CONTRIBUTIONS**

DJ conceived and designed the study. DJ and RR performed the literature search and selection process. RR, JB, and DJ performed the critical appraisal. DJ, CB, and RR analyzed and processed the included observational studies. DJ and CN analyzed and processed the included intervention studies. DJ wrote most of the paper with substantial contributions of CB, CN, and RR. FM continually supervised the paper and critically revised the full manuscript. All authors have read and approved the submitted version.

**FUNDING**

This work was supported by the German Research Foundation (DFG) and the Technical University of
Melina Schnitzius (Technical University of Munich) for her writing assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2019.00070/full#supplementary-material


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.