Low, WR, SANDERCOCK, GRH, FREEMAN, P, WINTER, M, Butt, J and Maynard, I

Pressure Training for Performance Domains: A Meta-Analysis.

Pressure Training for Performance Domains: A Meta-Analysis

William R. Low[1], G.R.H. Sandercock[1], Paul Freeman[1], Marie E. Winter[1], Joanne Butt[2], & Ian Maynard[1]

[1] University of Essex, UK

[2] Sheffield Hallam University, UK

Author note:

All correspondence should be addressed to William Low, School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. E-mail: wl18596@essex.ac.uk

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Studies have tested pressure training (PT) interventions in which performers practice physical or technical skills under simulated psychological pressure, but research has not yet synthesized the results of these studies. This meta-analysis assessed the magnitude of PT's effect on performance in sport and other high-pressure domains (e.g., law enforcement). A secondary purpose was to investigate how domain, dose, experience, and the type of task moderated the effectiveness of interventions. A study was included if it was peer-reviewed, conducted a PT intervention for sport or another high-pressure domain, and quantitatively compared a PT group to a control group on posttests under pressure. Fourteen studies in sport ($k = 10$) and law enforcement ($k = 4$) were included. Participants ($n = 394$) were novices, semi-professional athletes, elite athletes, and police officers. After removal of an outlier, the mean effect was medium ($g = 0.67$, 95% CI [0.43, 1.12]) with low heterogeneity ($I^2 = 17.1\%$). Subgroup analysis did not indicate clear moderators of performance but did reinforce that PT can benefit both novice and experienced participants on open and closed tasks across different domains. The results suggest coaches and instructors should create pressurized training environments rather than relying on greater amounts of training to help performers adjust to pressure. Future research should develop practical pressure manipulations, conduct retention tests, and measure performance in competitive or real-life scenarios.

Keywords: stress inoculation, stress exposure, sport, law enforcement, performance under pressure, meta-analysis, systematic review

Pressure Training for Performance Domains: A Meta-Analysis

The adages “practice how you play” or “train as you fight” demonstrate that domains such as sport and military understand that training should replicate performance as closely as possible to improve performance. Defined as “any factors or combination of factors that increase the importance of performing well on a particular occasion” (Baumeister, 1984, p. 610), psychological pressure is inherent to sport and other high-pressure domains, such as law enforcement (Hanton, Fletcher, & Coughlan, 2005; Nieuwenhuys & Oudejans, 2011). Research has studied whether training under pressure improves performance under pressure (e.g., Bell, Hardy, & Beattie, 2013). This pressure training (PT) is based on stress inoculation training (Meichenbaum, 2007) and involves physically practicing domain-specific skills under simulated pressure. Studies have also called PT “anxiety training” (e.g., Oudejans & Pijpers, 2009), “acclimatization training” (e.g., Beseler, Mesagno, Young, & Harvey, 2016), and “self-consciousness training” (e.g., Beilock & Carr, 2001). Despite their different names, these interventions all attempted to increase perceived pressure in training to enable participants to maintain or even improve performance under pressure.

PT can manipulate pressure by increasing either demands or consequences of a participant’s performance; however, delivering consequences seems to have a stronger effect upon anxiety than increasing demands does (Stoker et al., 2017). In sport, athletes can face loss of playing time, negative press, crowd derision or other consequences if they perform poorly. To simulate the pressure of these consequences, interventions have added monetary rewards (e.g., Oudejans & Pijpers, 2010), punishments (e.g., Bell et al., 2013), and perceived evaluation by coaches (e.g., Beseler et al., 2016). In other high-pressure domains, PT consequences can be inherent to the task and felt immediately (e.g., an antagonist firing back at police; Nieuwenhuys & Oudejans, 2011). PT may not perfectly replicate competition or life-threatening scenarios, but evidence suggests that anxiety in training can still help even if it is less severe than the anxiety felt during actual performance (Oudejans & Pijpers, 2010).

PT is distinct from other training methods that also manipulate conditions to prepare athletes and professionals for performance. For example, in a constraints-led approach to skill acquisition (Davids, Button, & Bennett, 2008), a soccer coach might train players' ball control by limiting the number of touches each player can take at a time. Like PT, this approach simulates performance conditions because players may not have the luxury of taking several touches in competition. However, PT and a constraints-led approach improve performance through different avenues: A constraints-led approach develops technical skills whereas PT trains the ability to cope with psychological pressure while performing those skills. Headrick, Renshaw, Davids, Pinder, and Araújo (2015) have acknowledged that training would better represent performance by incorporating emotional constraints experienced when performing. Pressure is one such constraint, and it can influence achievement in sport and safety in domains including medicine and law enforcement (Hardy et al., 2017; Arora et al., 2009; Vickers & Lewinski, 2012).

Although PT does not strictly teach physical or technical skills, it must combine the exposure to pressure with the simultaneous practice of such skills. For example, Oudejans and Pijpers (2009) found that dart players who practiced under pressure maintained subsequent performance in a pressurized posttest whereas performance declined for players who were merely exposed to pressure. PT does not just train the ability to cope with anxiety; instead, it trains the ability to cope while simultaneously executing skills or making decisions. PT is not necessarily a separate exercise from a performer's normal training regimen because a coach or instructor can increase pressure during an already-scheduled exercise. For instance, if a basketball team already practices free throws, then practicing free throws under pressure does not necessarily take much more time. Therefore, PT enhances existing training rather than introducing a completely new and unfamiliar exercise.

Systematic reviews have supported the effectiveness of PT (Gröpel & Mesagno, 2017; Kent, Devonport, Lane, Nicholls, & Friesen, 2018). In Kent et al. (2018), all five PT or "simulation training" interventions improved performance under pressure whereas all other interventions, such as cognitive-behavioral workshops and emotional regulation strategies, produced mixed results. In Gröpel and Mesagno's (2017) systematic review of choking interventions, eight out of nine PT studies ("acclimatisation training" or "self-consciousness training") led to statistically significant improvements in performance under pressure. Even though these findings are promising, they do not illustrate the magnitude of PT's effect on performance. Kent et al. (2018) acknowledged that a meta-analysis would have been inappropriate in their review because the variety of interventions and populations produced significant heterogeneity. Similarly, the mix of interventions in Gröpel and Mesagno (2017) may have also precluded meta-analysis. A review focused exclusively on PT interventions could have enough homogeneity to quantify their effect.

Comparing Kent et al. (2018) and Gröpel and Mesagno (2017) also reveals a need to more thoroughly assess PT research. These two reviews included only one of the same PT studies (i.e., Bell et al., 2013), and relevant literature could also include research on domains other than sport. Law enforcement and other domains inherently operate under pressure and already simulate their operating environments in training (e.g., Saus, Johnsen, Eid, Andersen, & Thayer, 2006). Systematic reviews in these domains have examined training of non-technical skills, such as teamwork (O'Dea, O'Connor, & Keogh, 2014), but no study has reviewed training for the domains' psychological pressures.

Sport does not have the same life-or-death risks associated with law enforcement, medicine, or aviation, but all of these domains require coping with pressure and have already learned from each other to improve training (Arora et al., 2009; Hanton et al., 2005). Medicine has adopted aviation's crew resource management training (Hamman, 2004; O'Dea et al., 2014) as well as athletes' cognitive training techniques, such as mental imagery (Wallace et al., 2017). Sport psychology has also informed military training (e.g., Fitzwater, Arthur, & Hardy, 2018). Despite the prevalence of pressure and the interest in improving training, little research has compared how these domains create and train in pressurized environments.

Even if PT has unique effects in sport compared to other domains, any differences could highlight the potential for learning across domains. Some heterogeneity is to be expected in a meta-analysis because included studies rarely all use the same methods and study the same participants (Higgins, 2008), and such heterogeneity would be expected especially for PT because these interventions can vary on several characteristics. Dose, or the number of PT sessions, has ranged from a single session (e.g., Beilock & Carr, 2001) to multiple sessions per week for several months (e.g., Bell et al., 2013). PT has been examined in novices and professionals (e.g., Liu, Mao, Zhao, & Huang, 2018; Oudejans, 2008), and PT can train performance of closed or open tasks under pressure (e.g., Alder, Ford, Causer, & Williams, 2016; Lewis & Linder, 1997). In closed tasks (e.g., golf putting), the performer chooses when to start executing a skill. In open tasks, the performer must execute a skill in response to a changing environment. Hitting a groundstroke in tennis is an open skill because the player must respond to the speed and location of an opponent's shot. Reviewing PT research could identify characteristics of PT associated with certain domains. Subgroup analysis could then quantify whether these characteristics moderated PT's effect, and results could provide rationale for one domain to adopt the best practices of another.

Findings of such a review could illustrate PT's value relative to other interventions and guide the timing, context, and design of PT. From a theoretical perspective, this synthesis could support or challenge potential explanations for PT's effects. Therefore, the current study's purpose was to assess the magnitude of PT's effect on performance under pressure in sport and other high-pressure domains. PT was defined as physically practicing domain-specific skills under simulated pressure. A secondary purpose was to explore if and how domain, dose, task type, and experience each moderated PT's effect.

Method

Literature Search

The method of this review followed PRISMA guidelines. Search terms were based on titles and keywords of PT studies already known to the authors, and six Boolean combinations were used to search MEDLINE, PsycINFO, PsycARTICLES, and SPORTDiscus. These databases were searched together in one search of EBSCOHost in August 2019. Boolean combinations were: 1) "pressure training" OR "practice with anxiety" OR "acclimatization training" OR "resilience training", 2) performance under pressure AND sport AND training, 3) "practice under pressure" OR "performance under pressure" OR "anxiety training" OR "acclimatization training," 4) performance under pressure AND anxiety AND training, 5) (simulation training or simulation education or simulation learning) AND anxiety, and 6) ("stress exposure training" or "stress inoculation training" or "stress training") AND performance. Searches were limited to scholarly journals, and they were not limited to any particular dates because this review was the first to examine PT exclusively.

Figure 1 illustrates the search and sifting process. The first and fourth authors independently sifted the search results by title and abstract, compared results, and resolved disagreements through discussion. Full text was examined when titles and abstracts were insufficient to determine eligibility. The first author also conducted backward and forward reference searching of studies after the final set of included studies from the search was determined. For the backward search, reference lists of these studies were scanned for other eligible studies. For the forward search, the "cited by" functions in the databases SCOPUS, Web of Science, and Google Scholar were used to identify articles that have since cited any of the already-included studies. Results were sifted by title, abstract, and full.

Inclusion Criteria

Studies were included if they: 1) trained and tested individuals on domain-specific skills, 2) conducted an intervention in which participants physically trained under simulated pressure, 3) compared an experimental group with a control group in a randomized or non-randomized study, 4) quantitatively measured each group's performance outcomes in a high-pressure posttest, 5) were written in English, and 6) were peer-reviewed and empirical. Inclusion was not limited to participants' level of experience because subgroup analysis was determined *a priori* to analyze level of experience. The fourth criterion specified performance in posttests because few sport psychology studies have measured performance in actual competition or real-life scenarios (Martin, Vause, & Schwartzman, 2005).

Data Items and Collection

The following pre-determined information was collected from each included study: 1) experimental design, 2) total *n*, 3) domain, 4) experience, 5) task, 6) task type (open or closed), 7) dose, and 8) pressure manipulations. According to the framework developed by Stoker, Lindsay, Butt, Bawden, and Maynard (2016), pressure manipulations were classified as forfeits (e.g., cleaning a changing room; Bell et al., 2013), rewards (e.g., money), judgment (e.g., evaluation by coaches), task stressors (e.g., time to complete a task), performer stressors (e.g., fatigue), or environmental stressors (e.g., noise). The first author completed a coding sheet with each variable for each study, and the fourth author verified the data. Six disagreements were resolved through discussion.

Mean posttest scores and standard deviations were extracted from articles or obtained by e-mailing authors. Four authors were e-mailed, and two responded with the requested data. GetData Graph Digitizer (http://getdata-graph-digitizer.com) was used to estimate data from graphs when means could not be obtained from articles or contact with authors. Standard errors and sample sizes were used to calculate standard deviations for each group for studies that did not report standard deviations.

Assessment of Bias

Risk of bias in randomized studies was assessed using the Cochrane Collaboration's tool for assessing risk of bias (Higgins & Green, 2011). For each study, the first and fourth authors assessed risks of selection, performance, detection, and attrition biases as low, high, or unclear. The authors evaluated non-randomized studies for the same biases using the Risk of Bias Assessment tool for Nonrandomized Studies (Kim et al., 2013). Studies that did not explicitly state if they were randomized were considered to be non-randomized.

It was anticipated that most studies would share unclear or high risks for many categories of bias because psychological studies do not typically follow procedures such as allocation concealment or blinding of researchers. Therefore, this assessment was intended to compare the included studies with each other and identify any bias that could distinguish studies within the review. For example, if risk of one bias was high in half the studies and low in the other half, then that bias would warrant further analysis to see if it affected results.

To assess bias across studies, a funnel plot displayed each study's effect size against the study's precision (i.e., standard error). Poor methodological designs or poor analysis can inflate effect sizes in small studies, and publication bias may prevent publication of studies with statistically non-significant results. Asymmetry in the funnel plot and a significant result from Egger's test would suggest the presence of publication bias or small-study effects.

Summary Measures and Planned Method of Analysis

The effect of PT was measured by the standardized mean difference (Hedges' *g*) between posttest performance scores of control and experimental groups. Each study was also inspected for differences between experimental and control groups at baseline. Hedges' *g* was used because it corrects for bias from small samples (Lakens, 2013). Using the DerSimonian and Laird approach in Stata, a random-effects model calculated an effect size and 95% confidence interval for each study as well as a pooled effect size and its 95% confidence interval. The heterogeneity of study characteristics supported a random-effects model, which assumes that all the studies represent different, but related, interventions (Higgins & Green, 2011). A random-effects model also allows inferences to generalize beyond included studies whereas results of fixed-effects models only apply to included studies (Field & Gillett, 2010). Effect sizes of 0.2, 0.5, and 0.8 were interpreted as small, medium, and large, respectively (Cohen, 1988). I^2 was calculated to measure heterogeneity. Expressed as a percentage, I^2 represents the variation across results due to heterogeneity among studies rather than chance (Higgins, Thompson, Deeks, & Altman, 2003).

Pre-specified additional analyses tested four potential moderators of PT effectiveness: domain, dose, experience, and task type. Domain referred to sport or another field (e.g., aviation, law enforcement, medicine) and was examined because differences in population, technical skills, and consequences of performance might influence PT's effectiveness. Dose referred to the number of PT sessions, and it was analyzed to help coaches and sport psychology practitioners determine how much PT they should conduct to improve performance. It would also guide future research because doses that are too short or too long could confound results of otherwise well-designed PT. Participants' experience in the domain being tested was examined because psychological interventions have had different effects for novices and experienced performers (e.g., Feltz & Landers, 1983). Many sports and occupations involve a mix of open and closed tasks, so task type was examined because the applicability of PT to each domain may depend on whether PT can improve performance on either type of task. A pooled Hedges' g, 95% confidence interval, and I^2 were calculated for each subgroup.

Five special circumstances required processing data to make them suitable for the meta-analysis. First, some performance measures (e.g., mean radial distance in golf putting; Beilock & Carr, 2001) were reversed so that greater values represented better performance, which aligned with measures in the other studies. Second, only two groups were compared even if a study had more than two groups (e.g., control, low-anxiety training, and high-anxiety training; Lawrence et al., 2014). Groups that physically trained under low pressure were used as the control group, instead of groups that did not train at all. Third, measures were averaged when a study had multiple continuous measures of performance (Bell et al., 2013). Fourth, performance was compared on posttests, rather than retention tests, because only one study conducted a retention test (Nieuwenhuys & Oudejans, 2011). Posttests assessed the effects of PT immediately after the intervention whereas a retention test would take place weeks or months after the intervention to assess how long effects were sustained. Finally, for studies that tested participants under low and high pressure (e.g., Oudejans & Pijpers, 2009), only scores from high-pressure posttests were used to calculate effect sizes.

Results

A total of fourteen studies were included in the meta-analysis. Ten studies were found in the database search. Four studies were found via backward searching. Zero studies were found via forward searching. Interrater agreement was 89% after reviewing titles, 97% after reviewing abstracts, and 92% after reviewing full texts. Case studies did not meet all inclusion criteria, but some case studies provided additional examples of PT interventions (Mace & Carroll, 1986; Mace, Eastman, & Carroll, 1986).

Study Characteristics

Table 1 illustrates characteristics of the included studies. Ten examined sport, and four examined law enforcement. Studies in any high-pressure domain were eligible for inclusion, but sport and law enforcement were the only ones with studies that met all the inclusion criteria. The included studies had a total of 394 participants and mean sample size of 28 participants (*SD* = 20). Participants were novices, trainees, semi-professionals, professionals, and international-level athletes. Doses ranged from 1 to 46 sessions of PT. Some studies used multiple pressure manipulations, and other studies used only one. Judgment was the most common ($k = 8$), followed by rewards ($k = 6$) and forfeits ($k = 4$).

Risk of Bias

Table 2 illustrates the results of the bias assessments. No single type of within-study bias distinguished studies into subgroups because there was little variation in their ratings on each category. Interrater agreement was 86%. A relatively symmetrical funnel plot and a non-significant Egger's test result ($P = 0.12$) showed no indication of significant publication bias or small-study effects across studies.

Mean Effect

The forest plot in Figure 2 presents the individual and pooled effect sizes, 95% confidence intervals, and the weight of each study. Across the included studies, PT had a large positive effect on performance under pressure for experimental groups when compared to control groups that did not receive PT ($g = 0.85$, 95% CI [0.37, 1.34]). Only Bell et al. (2013) had a significant difference between experimental and control groups at baseline on one performance measure, and this difference was balanced by no significant difference between groups on a second measure. Heterogeneity between studies was high ($I^2 = 78.4\%$).

The forest plot showed that one study (Liu et al., 2018) could be responsible for much of the high heterogeneity, so sensitivity analysis was conducted to measure the influence of each study on the mean effect. The mean effect was re-calculated while omitting each study one at a time. Omission of Liu et al. (2018) decreased Hedges' *g* from 0.85 to 0.67 and the upper limit of the 95% confidence interval from 1.33 to 0.94. In contrast, when any other study was omitted, Hedges' *g* was at least 0.83, and the upper limit of the 95% confidence interval was at least 1.34. Omission of Liu et al. (2018) also decreased I^2 from 78.4% to 17.1%. This more conservative estimate indicates a medium effect with a more precise 95% confidence interval ([0.41,0.94]).

Because of Liu et al. (2018)'s disproportional influence, it was omitted from the preplanned subgroup analyses. When heterogeneity is due to study characteristics, subgroup analysis can identify which characteristics are responsible, but high heterogeneity due to a single study would make results of subgroup analysis difficult to interpret. Thus, this omission made subgroup analysis of the remaining studies more robust.

Subgroup Analysis

Table 3 summarizes the effects of PT in each subgroup for the preplanned moderator variables: domain, dose, task type, and experience. Domain was coded as either "sport" or "law enforcement." Dose was coded as "short" (one PT session), "medium" (2-5 sessions), or "long" (over five sessions). Task type was either "open" or "closed." For experience, participants were divided into "novice" or "experienced" subgroups. All but one subgroup (long-dose interventions) had moderate effects, so none of these variables significantly moderated performance under pressure. For each variable, one subgroup's confidence interval encompassed the entire confidence interval of the other subgroup(s). This overlap suggests that little difference, if any, existed between PT's effects among subgroups. However, heterogeneity did distinguish subgroups and warrants interpreting similarities in effect size with caution. Long-dose interventions had the smallest effect of any subgroup ($g = 0.42$, 95% CI [-0.65, 1.50]) but also had the fewest studies ($k = 3$) and the highest heterogeneity ($I^2 = 73.1\%$). Although heterogeneity was only moderate among experienced participants (I^2 = 48.9%), it was lower for novices (I^2 = 0.0%). It should also be noted that all studies with novices overlapped with short-dose interventions.

Discussion

The main purpose of this meta-analysis was to assess the effectiveness of PT for enhancing performance under pressure. A secondary purpose was to explore if and how domain, dose, task type, and experience each moderated the magnitude and direction of PT's effect. Fourteen studies were included. Although studies from any high-pressure domain were eligible for inclusion, sport and law enforcement were the only domains represented. The range of the law enforcement studies was narrow: They all trained shooting skills, and three of the four studies were conducted by the same authors (Nieuwenhuys & Oudejans, 2011; Nieuwenhuys, Savelsbergh, & Oudejans, 2015; Oudejans, 2008). Studies have examined PT in firefighting and medicine (e.g., Baumann, Gohm, & Bonner, 2011; DeMaria et al., 2010), but they did not meet all inclusion criteria.

Results supported previous systematic reviews that found PT interventions consistently improved performance under pressure (Gröpel & Mesagno, 2017; Kent et al., 2018). Both previous reviews compared PT with other choking or coping interventions, but their reliance on statistical significance limited conclusions. Meta-analysis allowed the current review to measure the magnitude of PT's effect on performance under pressure. The included studies had a large positive effect (g = 0.85, 95% CI [0.37, 1.34]). This effect represents between-group differences on high-pressure posttests, so it suggests that performers who receive PT outperform others who do not receive PT. It does not, however, describe how that performance under high pressure compares to performance under low pressure. Included studies whose effect sizes were similar to this overall effect more concretely illustrate the meaning of the result. In Lawrence et al.'s (2014) experiment 1, the experimental group made more than 2.5 more putts than the control group did out of 25 total putts. In Nieuwenhuys and Oudejans (2011), police officers who received PT were 14 percent more accurate firing at an opponent than the control group was in the posttest.

After removal of an outlier with an especially large positive effect (Liu et al., 2018), the overall effect of PT was moderate (*g* = 0.67, 95% CI [0.41, 0.94]). Differences between the SWAT trainees in Liu et al. (2018) and novices in other studies could explain the large effect size. For example, the trainees may have been more motivated than other novices because the task was related to the trainees' careers.

This moedium effect of PT approximated the effects of other interventions for performance enhancement. It is within the 95% confidence interval of 0.22–0.92 (Hedges' *g*) that Brown and Fletcher (2017) found in their meta-analysis of various psychological and psychosocial interventions in sport, including pre-performance routines, self-talk, and imagery. Rather than competing with these interventions, PT may complement them in applied practice because PT could provide a more ecologically valid setting to practice routines, attentional training, or other techniques used during performance.

Bell et al. (2013) found PT was effective when combined with mental skills training; however, the remaining studies suggested PT alone can improve performance. According to Nieuwenhuys and Oudejans' (2017) model, pressure can prompt performers to increase mental effort as they become more concerned with performing well, and PT may train performers to direct this effort to completing their task rather than worrying about the pressure. Oudejans and Pijpers (2009) found that their control and experimental groups both increased effort in posttests under anxiety, but only the experimental groups' efforts improved performance. The two groups both remained anxious in posttests. Thus, rather than reducing anxiety, PT appeared to acclimatize participants to performing with anxiety.

PT effects were also consistent across domains. Police and athletes both performed better under pressure after PT. They did test under the same pressure manipulations used in their PT rather than real-life or competitive pressures (e.g., “soap” bullets instead of real bullets), which warrants more research to examine how well PT would translate to competition or an encounter with a suspect. The differences between control and experimental groups do imply that pressure can limit performance, so the results at least highlight the need to prepare for such pressure in both domains. One difference between the domains is that all police studies trained open tasks whereas most sport studies trained closed tasks. The open tasks were “extended” in that they involved a continuous series of opportunities to perform skills (e.g., firing multiple shots, reloading the weapon, and moving after each shot; Nieuwenhuys & Oudejans, 2011). Because many sports involve mostly extended open-task sequences, training these tasks in PT could prepare athletes for a wider variety of situations and train the ability to sustain that performance throughout a sequence.

Novices and experts both improved moderately after PT. The positive effect on experienced participants demonstrated that performers who are physically or technically skilled could still improve under pressure. Experience in one’s domain does not guarantee quality performance under pressure (e.g., Alder et al., 2016). For novices, improvements could be explained by the specificity of practice hypothesis, which suggests individuals perform better when they have learned under the same conditions in which they perform (e.g., high pressure; Cassell, Beattie, & Lawrence, 2018).

Interventions with five or more PT sessions had the smallest effect on performance under pressure. This finding contrasts recommendations in sport psychology for consistent, long-term interventions (Fifer, Henschen, Gould, & Ravizza, 2008), but the small number of these studies and their varied results (Table 3) show that more studies are needed to determine appropriate amounts of PT. Furthermore, we can speculate that results could differ if they were measured on retention tests because the advantage of long interventions could be in sustaining performance under pressure throughout a competitive season or career. Many of the scenarios simulated in PT studies (e.g., game-winning free throws) may only occur occasionally and unpredictably for each individual performer, so he or she may need to train under pressure consistently to stay prepared for such scenarios when they do occur.

Applied Implications

Because control groups physically practiced as much as experimental groups did, the between-group differences in performance should encourage leaders to increase pressure in practice, not just the amount of practice. Challenges help individuals develop psychological skills, and "constructed challenges," such as PT, develop these skills more intentionally than waiting for opportunities to occur naturally (Collins, Macnamara, & McCarthy, 2016, p.3). PT also contrasts approaches to learning that center around leaders or practitioners providing verbal explanations or demonstrations. While Bell et al. (2013) complemented PT with mental skills training, the remaining studies suggested that a practitioner would not have to explicitly teach mental skills for participants to acclimatize to pressure during PT. That is, participants seemed to adapt to pressure on their own. When preparing performers for pressure, leaders can create a pressurized atmosphere in which performers can independently learn to perform. This PT should take place in a facilitative environment in which leaders balance the challenge of pressure with support, such as strong coach-athlete relationships and encouragement to learn from mistakes (Fletcher & Sarkar, 2016).

Coaches or instructors could consider introducing appropriate amounts of pressure early in a learner's development. PT's effectiveness for novices illustrates that individuals might not have to master a skill before training it under pressure. Furthermore, when learners train while feeling emotions of competition, they may be more engaged and also discover the emotions, thoughts, and behavior that they need to perform optimally (Headrick et al., 2015).

Simulating such pressure may be more feasible if coaches and practitioners utilize stressors inherent to the task being trained. Despite increasing anxiety successfully, sport studies relied on external sources of pressure, including monetary rewards, that would be impractical for coaches to replicate regularly. Police, in contrast, faced consequences that were directly connected to their experimental task, such as shooting a live "hostage" (with a "soap" bullet) if they missed their target (Liu et al., 2018). These tasks also took place in simulated performance contexts, including realistic physical surroundings and verbal communication with suspects when first encountered (Nieuwenhuys et al., 2015). Similarly, situating PT in a simulated performance context could provide sources of pressure that are absent when individuals train a skill isolated from the flow of competition. For example, if basketball players pressure trained free throws during a practice game, or "scrimmage," during a training session, they would face stressors inherent to the scrimmage itself (e.g., failing to score easy points) as well as external stressors (e.g., judgment from coaches).

Future Directions & Limitations

A limitation of this review is that it did not evaluate the effectiveness of different pressure manipulations. Because many studies combined multiple stressors from different categories in Stoker et al.'s (2016) framework of pressure manipulations, subgroup analysis of each category was not possible. Stoker et al. (2017) previously examined athletes' perceptions of pressure from different manipulations, but future research should test which manipulations help improve performance most. In addition, low-cost and practical manipulations need to be developed so coaches and instructors can regularly implement PT.

A first step in developing these manipulations would be to identify high-pressure situations and the sources of their pressure. Although higher pressure is often associated with higher stakes, subjective appraisals of a situation as a challenge or threat can also moderate the effect of pressure (Seery, 2011). Factors such as the situation's unpredictability or novelty can in turn influence appraisals (Thatcher & Day, 2008). Many studies have examined sources of stress for athletes (e.g., Hanton et al., 2005), but few have examined the factors that increase pressure specifically during competition. Because leveraging other factors could increase pressure without increasing the size of rewards or severity of forfeits, these manipulations would make longer interventions more feasible.

More studies on longer interventions are needed to recommend how often to implement PT. Despite the appeal of "quick fix" solutions, sport psychology practitioners have emphasized that time and commitment are essential for psychological training to have lasting effects (Fifer et al., 2008). Still, most studies conducted fewer than five PT sessions and did not attempt to extend findings in laboratory or practice settings to competition or real-life scenarios. The number of sessions varied widely among the long interventions (Bell et al., 2013; Beseler et al., 2016; Oudejans & Pijpers, 2009), so it remains unclear how much PT is necessary for individuals to perform consistently better under pressure. PT may work by systematically desensitizing performers to pressure, which would require repeated exposure rather than a single session of PT. Therefore, future studies should implement PT over several weeks, months, or an entire season to determine both minimum *and* maximum amounts of PT. Guidelines for maximum amounts are important to establish in case longer doses diminish perceived pressure during PT. Longer studies would also provide chances to investigate how mental skills training might influence the efficacy and optimal dose of PT.

The subgroup analysis only tested how variables moderated performance on posttests, but more differences between interventions may emerge if effects are also evaluated on their sustainability over time. Only one study conducted a retention test (Nieuwenhuys & Oudejans, 2011), so more studies are needed to measure how long athletes remain acclimatized to pressure. Such retention tests could help identify amounts of PT that generate permanent learning without diminishing the effects of pressure manipulations.

Research could also test whether improvements under pressure transfer across skills within a sport or domain. Existing studies have measured PT effectiveness by testing the same skills that were practiced during PT, so it is still unknown whether performance gains illustrate a general or situation-specific ability to perform under pressure. If PT trains a general ability, then training one skill (e.g., tennis serves) under pressure could enhance other skills (e.g., groundstrokes) under pressure too. If it trains a skill-specific ability, then performers may need to pressure train many skills to prepare for the variety of situations that they could face. Transfer tests should therefore be conducted to examine how pressure-trained skills compare with skills not trained under pressure.

To truly assess transferability and sustainability, performance should also be measured in competition or real-life scenarios. Differences between practice and competition limits the generalizability of findings in one setting to the other, but few studies in sport psychology have assessed interventions by measuring performance in competitions (Martin et al., 2005). In the current review, Bell et al. (2013) did find that their experimental group outperformed the control group in competition, but they measured overall performance rather than performance in pressure situations. Although training under mild anxiety has prevented choking under higher anxiety in laboratory settings (Oudejans & Pijpers, 2010), studies are needed to support this finding in real-life or competitive performance situations.

Conclusion

Meta-analysis of 14 studies found PT improved performance under pressure for a wide range of participants and tasks in sport and law enforcement. The mean effect was medium after an outlier was excluded. Although more research should examine the role of mental skills training in enhancing PT, individuals seemed to learn independently to perform under pressure when given chances to practice under pressure. Interventions varied in their domain, dose, participants' experience, and task type, but no single characteristic increased or decreased PT's effectiveness. More clear moderators may emerge if studies examine the sustainability of PT's effect over time and transferability across domain-specific skills.

References

*Alder, D., Ford, P. R., Causer, J., & Williams, A. M. (2016). The effects of high- and low-anxiety training on the anticipation judgments of elite performers. *Journal of Sport and Exercise Psychology*, *38*, 93–104. https://doi.org/10.1123/jsep.2015-0145

Arora, S., Sevdalis, N., Nestel, D., Tierney, T., Woloshynowych, M., & Kneebone, R. (2009). Managing intraoperative stress: What do surgeons want from a crisis training program? *American Journal of Surgery*, *197*, 537–543. doi: 10.1016/j.amjsurg.2008.02.009

Baumann, M. R., Gohm, C. L., & Bonner, B. L. (2011). Phased training for high-reliability occupations: Live-fire exercises for civilian firefighters. *Human Factors*, *53*, 548–557. https://doi.org/10.1177/0018720811418224

Baumeister, R. F. (1984). Choking under pressure : Self-Consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*, 610–620.

*Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, *130*, 701–725. https://doi.org/10.1037/0096-3445.130.4.701

*Bell, J. J., Hardy, L., & Beattie, S. (2013). Enhancing mental toughness and performance under pressure in elite young cricketers: A 2-year longitudinal intervention. *Sport, Exercise, and Performance Psychology*, *2*, 281–297. https://doi.org/10.1037/spy0000010

*Beseler, B., Mesagno, C., Young, W., & Harvey, J. (2016). Igniting the pressure acclimatization training debate: Contradictory pilot-study evidence from Australian football. *Journal of Sport Behavior*, *39*, 22–38.

Brown, D. J., & Fletcher, D. (2017). Effects of psychological and psychosocial interventions on sport performance: A meta-analysis. *Sports Medicine*, *47*, 77–99. https://doi.org/10.1007/s40279-016-0552-7

Cassell, V. E., Beattie, S. J., & Lawrence, G. P. (2018). Changing performance pressure between training and competition influences action planning because of a reduction in the efficiency of action execution. *Anxiety, Stress and Coping*, *31*, 107–120. https://doi.org/10.1080/10615806.2017.1373389

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Lawrence Erlbaum.

Collins, D. J., Macnamara, A., & McCarthy, N. (2016). Putting the bumps in the rocky road: Optimizing the pathway to excellence. *Frontiers in Psychology*, *7*(SEP), 1–6. https://doi.org/10.3389/fpsyg.2016.01482

Davids, K., Button, C., & Bennett, S. (2008). *Dynamics of skill acquisition: A constraints-led approach*. Champaign, IL: Human Kinetics.

DeMaria, S., Bryson, E. O., Mooney, T. J., Silverstein, J. H., Reich, D. L., Bodian, C., & Levine, A. I. (2010). Adding emotional stressors to training in simulated cardiopulmonary arrest enhances participant performance. *Medical Education*, *44*, 1006–1015. https://doi.org/10.1111/j.1365-2923.2010.03775.x

Feltz, D. L., & Landers, D. M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *Journal of Sport Psychology*, *5*, 25–57.

Field, A., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, *63*, 665–694. https://doi.org/10.1348/000711010X502733

Fifer, A., Henschen, K., Gould, D., & Ravizza, K. (2008). What works when working with athletes. *Sport Psychologist*, *22*, 356–377. https://doi.org/10.1123/tsp.22.3.356

Fitzwater, J. P. J., Arthur, C. A., & Hardy, L. (2018). “The tough get tougher”: Mental skills training with elite military recruits. *Sport, Exercise, and Performance Psychology*, *7*, 93–107. https://doi.org/10.1037/spy0000101

Fletcher, D., & Sarkar, M. (2016). Mental fortitude training: An evidence-based approach to developing psychological resilience for sustained success. *Journal of Sport Psychology in Action*, *7*, 135–157. https://doi.org/10.1080/21520704.2016.1255496

Gröpel, P., & Mesagno, C. (2017). Choking interventions in sports: A systematic review. *International Review of Sport and Exercise Psychology*, *10*, 1–26. https://doi.org/10.1080/1750984x.2017.1408134

Hamman, W. R. (2004). The complexity of team training: What we have learned from aviation and its applications to medicine. *Quality and Safety in Health Care*, *13*, 72–79. https://doi.org/10.1136/qshc.2004.009910

Hanton, S., Fletcher, D., & Coughlan, G. (2005). Stress in elite sport performers: A comparative study of competitive and organizational stressors. *Journal of Sports Sciences*, *23*, 1129–1141. https://doi.org/10.1080/02640410500131480

Hardy, L., Barlow, M., Evans, L., Rees, T., Woodman, T., & Warr, C. (2017). Great British medalists. In *Progress in Brain Research* (1st ed., Vol. 232, pp. 1–119). Elsevier B.V. https://doi.org/10.1016/bs.pbr.2017.03.004

Headrick, J., Renshaw, I., Davids, K., Pinder, R. A., & Araújo, D. (2015). The dynamics of expertise acquisition in sport: The role of affective learning design. *Psychology of Sport and Exercise*, *16*, 83–90. https://doi.org/10.1016/j.psychsport.2014.08.006

Higgins, J. P. T. (2008). Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology*, *37*, 1158–1160. https://doi.org/10.1093/ije/dyn204

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.). The Cochrane Collaboration. Retrieved from www.handbook.cochrane.org

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*, 557–560. doi: 10.1136/bmj.327.7414.557

Kent, S., Devonport, T. J., Lane, A. M., Nicholls, W., & Friesen, A. P. (2018). The effects of coping interventions on ability to perform under pressure. *Journal of Sports Science and Medicine*, *17*, 40–55. https://doi.org/10.1016/j.paid.2017.06.021

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H.-J. J., Sheen, S.-S. S., Hahn, S., … Son, H.-J. J. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*, *66*, 408–414. https://doi.org/10.1016/j.jclinepi.2012.09.016

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

*Lawrence, G. P., Cassell, V. E., Beattie, S., Woodman, T., Khan, M. A., Hardy, L., … Gottwald, V. M. (2014). Practice with anxiety improves performance, but only when anxious: Evidence for the specificity of practice hypothesis. *Psychological Research*, *78*, 634–650. https://doi.org/10.1007/s00426-013-0521-9

*Lewis, B. P., & Linder, D. E. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin*, *23*, 937–944. https://doi.org/10.1177/0146167297239003

*Liu, Y., Mao, L., Zhao, Y., & Huang, Y. (2018). Impact of a simulated stress training program on the tactical shooting performance of SWAT trainees. *Research Quarterly for Exercise and Sport*, *89*, 482–489. https://doi.org/10.1080/02701367.2018.1526003

Mace, R., & Carroll, D. (1986). Stress inoculation training to control anxiety in sport: two case studies in squash. *British Journal of Sports Medicine*, *20*, 115–117. https://doi.org/10.1136/bjsm.20.3.115

Mace, R., Eastman, C., & Carroll, D. (1986). Stress inoculation training: A case study in gymnastics. *British Journal of Sports Medicine*, *20*, 1399–141. doi: 10.1136/bjsm.20.3.139

Martin, G. L., Vause, T., & Schwartzman, L. (2005). Experimental studies of psychological interventions with athletes in competitions: Why so few? *Behavior Modification*, *29*, 616–641. https://doi.org/10.1177/0145445503259394

Meichenbaum, D. (2007). Stress inoculation training: A preventative and treatment approach. In P. M. Lehrer, R. L. Woolfolk, & W. E. Sime (Eds.), *Principles and practice of stress management* (3rd ed.). New York: The Guilford Press.

Nieuwenhuys, A., & Oudejans, R. R. (2017). Anxiety and performance: Perceptual-motor behavior in high-pressure contexts. *Current Opinion in Psychology*, *16*, 28–33. https://doi.org/10.1016/j.copsyc.2017.03.019

*Nieuwenhuys, A., & Oudejans, R. R. D. (2011). Training with anxiety: Short-and long-term effects on police officers' shooting behavior under pressure. *Cognitive Processing*, *12*, 277–288. https://doi.org/10.1007/s10339-011-0396-x

*Nieuwenhuys, A., Savelsbergh, G. J. P., & Oudejans, R. R. D. (2015). Persistence of threat-induced errors in police officers' shooting decisions. *Applied Ergonomics*, *48*, 263–272. https://doi.org/10.1016/j.apergo.2014.12.006

O'Dea, A., O'Connor, P., & Keogh, I. (2014). A meta-analysis of the effectiveness of crew resource management training in acute care domains. *Postgraduate Medical Journal*, *90*, 699–708. https://doi.org/10.1136/postgradmedj-2014-132800

*Oudejans, R. R. D. (2008). Reality-based practice under pressure improves handgun shooting performance of police officers. *Ergonomics*, *51*, 261–273. https://doi.org/10.1080/00140130701577435

*Oudejans, R. R. D., & Pijpers, J. R. (2009). Training with anxiety has a positive effect on expert perceptual-motor performance under pressure. *Quarterly Journal of Experimental Psychology*, *62*, 1631–1647. https://doi.org/10.1080/17470210802557702

*Oudejans, R. R. D., & Pijpers, J. R. (2010). Training with mild anxiety may prevent choking under higher levels of anxiety. *Psychology of Sport and Exercise*, *11*, 44–50. https://doi.org/10.1016/j.psychsport.2009.05.002

Saus, E.-R., Johnsen, B. H., Eid, J., Riisem, P. K., Andersen, R., & Thayer, J. F. (2006). The effect of brief situational awareness training in a police shooting simulator: An experimental study and the Royal Norwegian Navy. *Military Psychology*, *18*, 3–21. https://doi.org/10.1016/j.chb.2012.02.009

Seery, M. D. (2011). Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. *Neuroscience and Biobehavioral Reviews*, *35*, 1603–1610. https://doi.org/10.1016/j.neubiorev.2011.03.003

Stoker, M., Lindsay, P., Butt, J., Bawden, M., & Maynard, I. W. (2016). Elite coaches' experiences of creating pressure training environments. *International Journal of Sport Psychology*, *47*, 262–281. https://doi.org/10.7352/IJSP2016.47.262

Stoker, M., Maynard, I., Butt, J., Hays, K., Lindsay, P., & Adams Norenberg, D. (2017). The effect of manipulating training demands and consequences on experiences of pressure in elite netball. *Journal of Applied Sport Psychology*, *29*, 434–448. https://doi.org/10.1080/10413200.2017.1298166

Thatcher, J., & Day, M. C. (2008). Re-appraising stress appraisals: The underlying properties of stress in sport. *Psychology of Sport and Exercise*, *9*, 318–335. https://doi.org/10.1016/j.psychsport.2007.04.005

Vickers, J. N., & Lewinski, W. (2012). Performing under pressure: Gaze control, decision making and shooting performance of elite and rookie police officers. *Human Movement Science*, *31*, 101–117. https://doi.org/10.1016/j.humov.2011.04.004

Wallace, L., Raison, N., Ghumman, F., Moran, A., Dasgupta, P., & Ahmed, K. (2017). Cognitive training: How can it be adapted for surgical education? *Surgeon*, *15*, 231–239. https://doi.org/10.1016/j.surge.2016.08.003

Table 1

Characteristics of Studies Included in Meta-Analysis

Study	Design	*N*	Domain	Experience	Task	Task Type	Dose	Pressure Manipulation
Alder, Ford, Causer, and Williams (2016)	R	20	Badminton	International	Reading location of opponent serves	Open	3	Judgment
Beilock and Carr (2001): experiment 3	R	36	Golf	Novice	Putting	Closed	1	Judgment
Bell, Hardy, and Beattie (2013)	NR	41	Cricket	Elite youth	Batting against pace and batting against spin	Open	46	Forfeit
Beseler, Mesagno, Young, and Harvey (2016)	R	12	Australian football	Semi-professional	Set shots	Closed	14	Environmental, judgment, reward
Lawrence et al. (2014): experiment 1	R	16	Golf	Novice	Putting	Closed	1	Judgment, reward
Lawrence et al. (2014): experiment 2	R	16	Rock climbing	Novice	Horizontal indoor climbing	Closed	1	Judgment, reward
Lewis and Linder (1997)	NR	30	Golf	Novice	Putting	Closed	1	Judgment, reward
Liu, Mao, Zhao, and Huang (2018)	R	92	SWAT team	In training	Shooting in hostage rescue	Open	3	Environmental
Nieuwenhuys and Oudejans (2011)	R	27	Police	Experienced professionals	Handgun shooting	Open	4	Forfeit
Nieuwenhuys, Savelsbergh, and Oudejans (2015)	NR	34	Police	Experienced professionals	Shoot/don't-shoot decisions	Open	3	Forfeit
Oudejans (2008)	NR	17	Police	Experienced professionals	Handgun shooting	Open	3	Forfeit
Oudejans and Pijpers (2009): experiment 1	NR	17	Basketball	"Expert"	Free throws	Closed	9	Judgment, reward
Oudejans and Pijpers (2009): experiment 2	NR	17	Darts	"Experienced"	Dart throwing	Closed	1	Environmental
Oudejans and Pijpers (2010)	R	24	Darts	Novice	Dart throwing	Closed	1	Judgment, reward

Note. R = randomized; NR = non-randomized; *N* = total number of participants in control and experimental groups included in the meta-analysis.

Table 2

Risk of bias assessments results

Randomized studies							
Study	Selection: randomization	Selection: allocation	Performance	Detection	Attrition	Reporting	Other
Alder et al. (2016)	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Low
Beilock & Carr (2001)	Unclear	Unclear	Unclear	Unclear	Unclear	Low	High
Beseler et al. (2016)	Unclear	Unclear	High	Unclear	High	High	Low
Lawrence et al. (2014): expt. 1	Unclear	Unclear	Unclear	Unclear	Unclear	Low	High
Lawrence et al. (2014): expt. 2	Unclear	Unclear	Low	Low	Unclear	Low	High
Liu et al. (2018)	Unclear	Unclear	Unclear	Low	High	Low	Low
Nieuwenhuys & Oudejans (2011)	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Low
Oudejans & Pijpers (2010)	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Low

Non-randomized studies						
Study	Selection	Confounds	Measurement Exposure	Blinding	Incomplete Data	Selective Reporting
Bell et al. (2013)	Low	Low	Low	Unclear	High	Low
Lewis & Linder (1997)	Unclear	Unclear	Low	Low	Low	Low
Nieuwenhuys et al. (2015)	Unclear	Unclear	Low	Low	Unclear	Low
Oudejans (2008)	Unclear	Unclear	Low	Low	Unclear	Low
Oudejans & Pijpers (2009): expt. 1	High	Low	Low	Low	Unclear	Low
Oudejans & Pijpers (2009): expt. 2	Unclear	Low	Low	Low	Unclear	Low

Table 3
Effect of Moderator Variables

Moderator	Subgroup	*k*	*N*	*g*	95% CI	Effect descriptor	*P*	Within-group I^2 (%)
Domain	Sport	10	224	0.72	[0.45, 1.00]	Moderate	< 0.001	0.0
	Law enforcement	3	78	0.63	[-0.14, 1.39]	Moderate	0.107	60.5
Experience	Experienced	8	180	0.61	[0.17, 1.05]	Moderate	0.007	48.9
	Novice	5	122	0.77	[0.40, 1.14]	Moderate	< 0.001	0.0
Dose	Short	6	139	0.73	[0.38, 1.08]	Moderate	< 0.001	0.0
	Medium	4	98	0.72	[0.11, 1.33]	Moderate	0.021	51.3
	Long	3	65	0.42	[-0.65, 1.50]	Small	0.440	73.1
Task Type	Open	5	134	0.74	[0.27, 1.20]	Moderate	0.002	38.2
	Closed	8	168	0.65	[0.30, 0.99]	Moderate	< 0.001	12.2

Note. *k* = number of studies; *N* = total number of participants; *g* = Hedges' *g*; CI = confidence interval

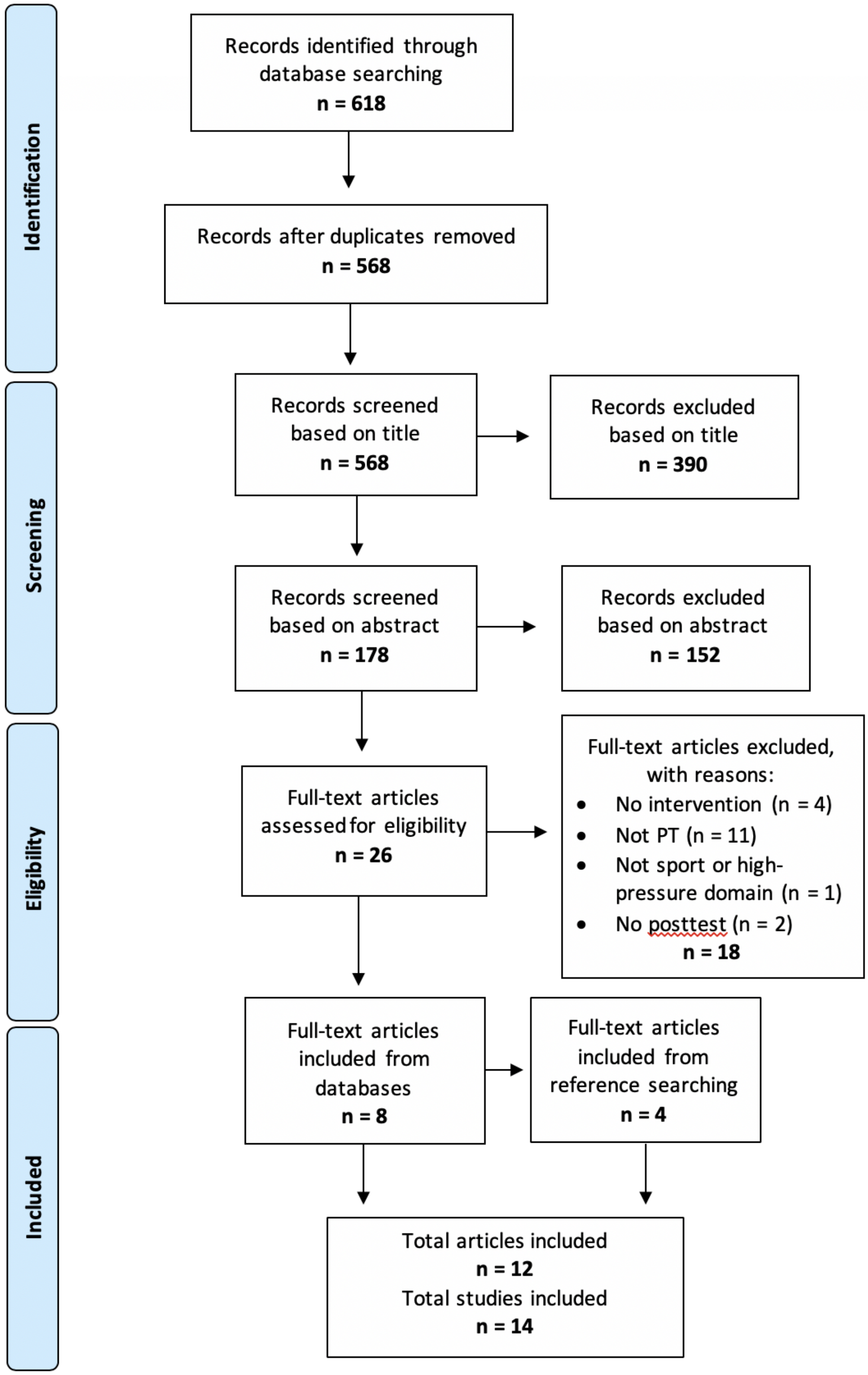

Figure 1. Identification of studies included in meta-analysis.

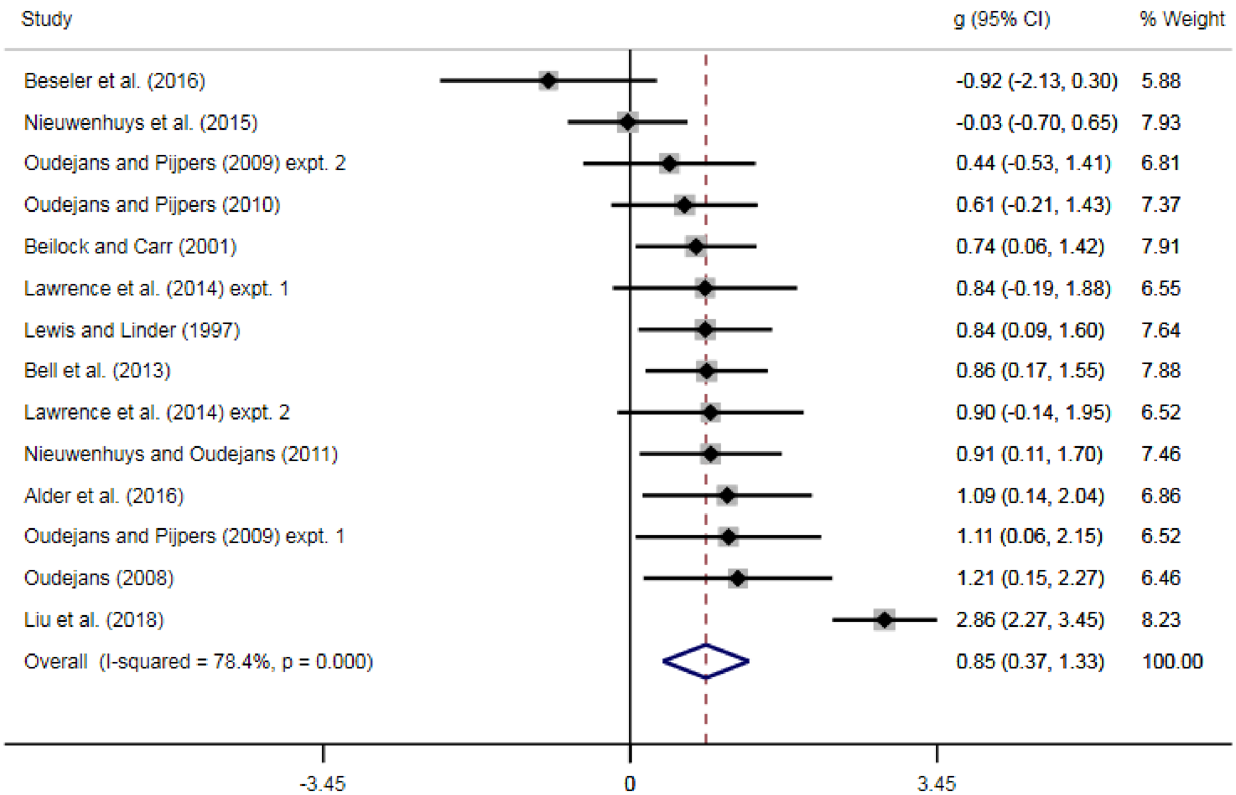

Figure 2. Forest plot of study effect sizes in ascending order.