Flow states in exercise: A systematic review

Patricia C. Jackman [a*], Rebecca M. Hawkins [a], Lee Crust [a] and Christian Swann [b, c]

[a] School of Sport and Exercise Science, University of Lincoln, Brayford Pool, Lincoln, United Kingdom

[b] School of Health and Human Sciences, Southern Cross University, Coffs Harbour, Australia

[c] Centre for Athlete Development, Experience and Performance, Southern Cross University, Coffs Harbour, Australia

* Corresponding author: Patricia Jackman, School of Sport and Exercise Science, University of Lincoln, Brayford Pool, Lincoln LN6 7TS, United Kingdom. Email: pjackman@lincoln.ac.uk

Highlights

- The first systematic review of flow research in exercise was conducted.
- Exergame design features, music, and virtual stimuli can affect at least some flow dimensions.
- A range of conceptual and methodological issues must be addressed in the field.
- Initial qualitative findings offer a potential step towards explaining the occurrence of flow.
- Researchers need to direct more attention towards developing an explanatory theory for flow states in exercise.

Abstract

Objectives: The purpose of this study was to systematically identify, review, synthesise, and appraise current literature on flow in exercise. By doing so, this study aimed to highlight gaps and future research directions that will help to advance understanding and application of flow states in this setting.

Design: A systematic review using PRISMA guidelines.

Methods: Eight electronic databases were searched in February 2019. Inclusion criteria were peer-reviewed studies focused on the investigation of flow in exercise. Exclusion criteria were studies that did not exclusively include exercise participants, or that focused on instrument development and/or validation. Data from included studies were extracted and reported in a narrative synthesis.

Results: A total of 26 studies that were conducted with 4478 participants met the inclusion criteria. Several issues with the conceptualisation and measurement of flow in exercise were identified, which makes it difficult to draw meaningful conclusions about this literature. Nevertheless, there is tentative evidence that exergame design features, music, and virtual stimuli can affect at least some dimensions of flow. While little attention has been directed towards developing an explanatory theory, initial findings concerning the contexts and process underlying flow occurrence could offer a potential avenue for progress.

Conclusions: The review advances knowledge by synthesising quantitative and qualitative evidence on flow states in exercise. By doing so, the review also highlights a range of conceptual and methodological issues in the field. Recommendations to address these issues and suggestions for making meaningful progress to develop understanding of flow states in exercise are advanced.

Keywords: adherence; clutch state; exergame; fitness; optimal experience; physical activity.

Flow states in exercise: A systematic review

Understanding optimal experiences is a key area of research in the field of positive psychology (Seligman & Csikszentmihalyi, 2000). The most advanced framework for understanding optimal experience is flow (Csikszentmihalyi, 1975), which is defined as an intrinsically rewarding psychological state in which individuals experience total absorption in an activity, perceptions of control, and a sense that everything clicks together, even in challenging situations (Csikszentmihalyi, 2002). Flow is most commonly conceptualised in terms of the nine-dimensions framework (Csikszentmihalyi, 2002), which describes the experience as an amalgam of: challenge-skills balance; action-awareness merging; clear goals; unambiguous feedback; concentration on the task at hand; sense of control; loss of self-consciousness; transformation of time; and autotelic experience. This multifaceted experience could be particularly relevant for exercisers, as flow has been associated with increased motivation (Schüler & Brunner, 2009) and long-term adherence (Elbe, Barene, Strahler, Krustrup, & Holtermann, 2016). In turn, flow can offer a potential foundation for building long-term exercise[1] engagement (Petosa & Holtz, 2013). Therefore, a robust evidence base is essential to inform the development of recommendations for exercisers and practitioners to increase the frequency and intensity of flow states in exercise settings.

Since the initiation of flow research in sport[2] in the early 1990's (e.g., Jackson, 1992), there has been extensive progress in understanding of flow in athletes (see Swann, Keegan, Piggott, & Crust, 2012 for a systematic review). While sport shares some characteristics with exercise (e.g., bodily movement), these subcategories of physical activity are considered to be distinct (Caspersen, Powell, & Christenson, 1985; World Health Organisation, 2018). Indeed,

[1] Exercise is defined as a structured, planned, repetitive bodily movement undertaken to enhance or maintain at least one aspect of physical fitness (Caspersen, Powell, & Christenson, 1985; World Health Organisation, 2018).
[2] Sport is defined as "an activity involving physical exertion, skill and/or hand-eye coordination as the primary focus of the activity, with elements of competition where rules and patterns of behaviour governing the activity exist formally through organizations" (World Health Organisation, 2018, p. 101).

researchers have highlighted the need to view exercise-specific literature on flow independently on the basis that findings in sport might not be transferable to exercise (e.g., Grove & Lewis, 1996). In accordance with this distinction, a great deal of research has focused on developing understanding of flow in exercise settings, with growing literature on the experience across a range of exercise activities (e.g., Elbe et al., 2016; Karageorghis, Vlachopoulos, & Terry, 2000; Monedero, Murphy, & O'Connor, 2017).

Despite the considerable growth in flow research in exercise, no studies have yet reviewed this literature systematically. A systematic review was conducted on the experience, occurrence, and controllability of flow in elite sport (Swann et al., 2012), while various book chapters and narrative reviews have been published on flow in sport and exercise (e.g., Kimiecik & Harris, 1996; Jackson & Kimieick, 2008; Swann, Piggott, Schweickle, & Vella, 2018). However, the increasing body of literature specifically on flow in exercise has yet to be systematically reviewed and synthesised. To move forward with flow research in exercise, it is important and timely to systematically review the knowledge base to provide a comprehensive, unbiased account of current understanding in this domain. Further, given that systematic reviews can highlight research gaps (Petticrew & Roberts, 2006) and provide direction for further empirical studies (Gurevitch, Koricheva, Nakagawa, & Stewart, 2018), a review of this nature could provide a robust foundation of knowledge to inform the design and direction of future research that seeks to advance understanding of flow in exercise.

Therefore, the purpose of this study was to systematically review, synthesise, and appraise existing research on flow in exercise. Specifically, this review aimed to address the following questions: (i) how has flow been measured in exercise?; (ii) how has flow been conceptualised in exercise?; (iii) what is the quality of research in this field?; (iv) what evidence is available regarding mechanisms of occurrence and strategies for inducing flow?; and (v) what are the correlates and outcomes of flow in exercise? By doing so, this study sought to provide a platform for future research on flow in exercise, with the ultimate goal of helping exercisers experience flow states more regularly and reliably.

Method

Protocol

This systematic review followed guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement (see Supplementary File 1 for checklist; Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009).

Search Strategy

The online search was conducted using eight electronic databases, comprising: Academic Search Complete; MEDLINE; PsycARTICLES; PsycINFO; PubMed; Scopus; SPORTSDiscus with Full Text; and Web of Science (Core Collection). Databases were searched four times (July 2018 - February 2019) to ensure that all newly published or in press articles were included. The final search was conducted on February 1st 2019. The search string employed the following Boolean search terms: [flow AND (state* OR experienc* OR theory)] AND (exer* OR physical* activ* OR health* OR fit* OR gym*) AND (psycholog*) NOT (expirat* flow OR optic* flow OR *water flow OR blood flow OR gene* flow). These keywords were selected following: initial scoping searches by the first and fourth authors on the EBSCOhost database; critical discussions between both authors after the initial searches; and a review of the search strategy in a previous systematic review in a similar but distinct domain (Swann et al., 2012). The search string was adapted appropriately for each database and, where possible, results were limited to peer-reviewed journal articles in the English language. The first two blocks in the search string were searched in the title and abstract fields, while the second two blocks were searched, at a minimum, in the title, abstract, and keyword fields (see Supplementary File 2 for the full search strategy). A range of synonyms for exercise were employed in the second block to ensure the comprehensiveness of this review (cf. Siddaway, Wood, & Hedges, 2019). Truncated terms were used, where possible, to ensure that as many variant spellings of a term as possible were captured. All retrieved titles and abstracts were exported into Endnote X8 reference management software. Duplicates were identified through the automatic de-duplication feature and manual searches.

Eligibility Criteria

Inclusion and exclusion criteria were established to ensure that the review parameters were clearly defined and that all literature relevant to the review objectives was identified (Smith, 2018). The inclusion criteria (a-e) stipulated that included studies were required to: (a) be a peer-reviewed journal article in the English language; (b) contain original empirical data; (c) be published or in press prior to February 1st 2019 (when the final search was undertaken); (d) refer to flow (i.e., psychological state) and an activity defined as exercise, or matching the definition of exercise[1] (Caspersen et al., 1985; World Health Organisation, 2018), either in the title or abstract (i.e., if the study referred to sport, or other terms, and could not clearly be interpreted as exercise, it was excluded); and (e) refer to the investigation of flow in exercise as a purpose of the inquiry. The exclusion criteria (f-g) specified that studies were ineligible if they: (f) involved participants in activities[3] that were not relevant to the objectives of the review; or (g) focused on instrument development and/or validation.

Screening Process

All titles and abstracts were initially read and checked for eligibility against the aforementioned criteria by the first and second authors independently. A meeting then took place to discuss the results of the screening process and resolve any discrepancies. Full texts were obtained for studies that referred to sport and exercise to check whether data for

---

[3] Common activities that did not correspond with the definition of exercise[1] and were excluded from the review included: adventure or sport tourism; physical education; recreational physical activity (e.g., housework); and sport[2]. These activities were excluded on the basis that the motive for participation might not be improving physical fitness (e.g., individuals in physical education classes might be participating to meet educational requirements), which is necessary for an activity to be deemed as exercise[1].

participants in exercise could be separated (i.e., inclusion/exclusion criteria f; samples that mixed sport and exercise participants but could not be separated at the full text stage were excluded). After screening the titles and abstracts for compliance with the inclusion criteria, manual searching was performed by the first author to identify further potentially relevant articles through: (i) reference lists of each identified study; (ii) titles of publications by authors of each identified study; (iii) titles in journals that published the identified studies; and (iv) titles of forward citations of the identified studies on Google Scholar. This process was then repeated for all articles identified through manual searching until no further potentially relevant studies were located. After concluding the manual search, all full texts of the relevant articles were obtained and searched in detail for compliance with the eligibility criteria by the first and second authors independently. Both researchers met to discuss the outcomes of the screening process and agree reasons for excluding rejected studies. Any disagreements or uncertainties that arose were assessed by the fourth author and a final consensus was reached. To ensure the accuracy of the study selection process, the third author acted as a 'critical friend' (Smith & McGannon, 2018) by judging a random selection of 10 titles and abstracts, and five full texts against the eligibility criteria. The third author discussed their decisions with the first author, with both authors agreeing on the decision for each randomly selected article.

Data Extraction, Analysis, and Synthesis

Following the identification of all relevant studies, the first author read each included study twice to enhance familiarity with the data before extracting and synthesising findings (cf. Smith, 2018). For this review, the following data were extracted: authors; date of publication; sample characteristics; exercise activity/activities; design; measurement of dispositional flow (i.e., tendency to experience flow in general) or state flow (i.e., flow experience in a specific activity); method and procedure; and key findings (i.e., findings not relevant to the aim of the review were not extracted). A weighted mean age for the sample was calculated by multiplying the mean age by the number of participants that reported mean age for each study, with these values then added together and divided by the total quantity of participants that reported mean age in all included studies. Where possible, effect sizes (Cohen's *d*) were calculated for differences between means (see Supplementary File 3 for all effect sizes) using Comprehensive Meta-Analysis (Version 3; Borenstein, Hedges, Higgins, & Rothstein, 2015). Effect sizes were calculated based on means, standard deviations, and sample sizes. If such information was not available, *F*-statistics were used. WebPlotDigitizer (Rohatgi, 2019) was used to extract data that was only published in figures and could not be retrieved. Given the heterogeneity in study design and outcomes assessed, a meta-analysis was not performed to summarise the quantitative findings. Key findings from the included studies were reviewed and integrated through narrative synthesis (Popay et al., 2006), similar to previous systematic reviews (e.g., Norris, Didymus, & Kaiseler, 2017; Swann et al., 2012). This interpretative, iterative, and integrative process enabled a synthesis of the findings into five broad categories (see *Results*) and facilitated the construction of a textual summary.

Assessing Study Quality

The assessment of study quality helps to ensure that studies included in a systematic review reach an acceptable scientific standard (Smith, 2018). Study quality was assessed using the 16-item quality assessment tool (QATSDD; Sirriyeh, Lawton, Gardner, & Armitage, 2012), which can be used to assess the quality of quantitative and qualitative studies. The QATSDD contains a list of criteria for quantitative and qualitative studies that are rated on a 4-point scale, ranging from 0 (not at all) to 3 (complete). All studies were assessed with respect to relevant criteria (i.e., criteria that applied to quantitative, qualitative, or mixed method) with the exception of criterion 14, which was omitted due to recent criticism of reliability strategies for qualitative research (Smith & McGannon, 2018). The total score for each included study was divided by the maximum possible score and reported as a percentage (*M* = 63.57%) for standardisation purposes (see Table 1 and Supplementary File 4). In line with recommendations (Milner, 2015), study quality was assessed by multiple authors. Initial scoring of 25 studies was undertaken by the first author, and both the third and fourth authors each assessed half of the studies included in the review (i.e., 13 studies each). One paper (Swann, Jackman, Schweickle, & Vella, 2019) was not assessed by the first or fourth author due to potential conflict of interest, but was reviewed by the second and third authors. After this process, the first author collated and synthesised the scores. The inter-rater agreement coefficient indicated a substantial level of agreement ($\kappa = 0.75$). Any discrepancies were resolved through critical discussions between the first author and the research team in accordance with the critical friends process (Smith & McGannon, 2018). During these critical discussions, the authors highlighted some concerns with the study quality scores as the QATSDD was unable to detect many of the conceptual and methodological issues identified by this review (see *Measurement of Flow in Exercise* and *Conceptualisation of Flow in Exercise*). Thus, although study quality scores are presented, we advise that caution should be taken when using these scores to judge study quality and suggest that any appraisal of the evidence quality should consider issues identified by this review.

Results

The electronic database generated a total of 2422 records, with 1172 unique records remaining after the removal of duplicates. After screening the titles and abstracts, 1141 articles were excluded. An additional 12 articles were identified through manual searches. Full texts for the remaining 43 records were reviewed for eligibility. Following this process, 26 articles met the inclusion criteria and were included (Figure 1). Most excluded studies involved sport participants (*n* = 10). In addition, data for specific samples in three studies (Butzer, Ahmed, & Khalsa, 2016; Elbe et al., 2016; Elbe, Strahlet, Krustrup, Wikman, &

Stelter, 2010) were excluded as the samples did not satisfy the eligibility criteria (i.e., participants were not involved in activities matching the definition of exercise). Similarly, qualitative data for two studies (Karageorghis & Jones, 2014; Loveday & Burgess, 2017) were excluded as the qualitative component in each study did not seek to investigate flow (i.e., did not satisfy eligibility criterion e). Findings were segregated and reviewed in terms of five categories, which are presented below: (i) study characteristics; (ii) measurement of flow; (iii) conceptualisation of flow; (iv) mechanisms of occurrence and strategies to induce flow; and (v) correlates and outcomes of flow. Within these categories, higher-order and lower-order thematic sections are used to explicate the findings. Where relevant, the type of exercise activity and study design is included in the narrative to contextualise the findings.

[INSERT FIGURE 1 ABOUT HERE]

Study Characteristics

A summary of characteristics of studies included in the review is presented in Table 1. The sample comprised: 24 quantitative papers; one mixed methods paper; and one qualitative paper. All included studies were published between 1996 and 2019. Of the 4478 participants, 2820 (62.97%) were female, 1369 (30.57%) were male, and gender was not specified for 289 (6.46%) participants. The weighted mean age of samples in the review that included details on participant mean ages was 36.51 years. The most common study design was cross-sectional (n = 11), followed by experimental (n = 10), longitudinal (n = 4), and exploratory investigation (n = 2). Flow was studied in activities including: exergaming[4] (n = 13; further information on the specific exergames is provided in Table 1); traditional gym-based (TGB) exercise (i.e., balance and weight training; n = 7); running (n = 5); treadmill walking/running (n = 3); cycling (n = 2); tai chi (n = 2); yoga (n = 2); "cardiovascular exercise" (n = 1; Ersöz

[4] An exergame is a video console game that incorporates exercise (e.g., Maddison et al., 2007). In this review, the terms 'exergame' and 'exergaming' will be used as a synonym for: active video games; body-movement controlled video games; mobile exergames; and virtual-reality based exercise games.

& Eklund, 2017, p. 95); circuit-training (*n* = 1); dance aerobics (*n* = 1); duathlon (*n* = 1); half ironman (*n* = 1); hiking (*n* = 1); indoor climbing (*n* = 1); surfing (*n* = 1); and Zumba (*n* = 1). Two studies involved participants in unspecified exercise activities (Ersöz, Altiparmak, & Hülya Aşçı, 2016; Ersöz & Eklund, 2017). Most studies investigated flow in specific exercise bouts (*n* = 22) while the remainder examined dispositional flow (i.e., frequency of flow experiences *in general*; *n* = 4).

[INSERT TABLE 1 ABOUT HERE]

Measurement of Flow in Exercise

This section reviews the ways in which flow has been measured in exercise to date. Findings are discussed in terms of: (i) quantitative research on flow in exercise; (ii) qualitative research on flow in exercise; and (iii) mixed-method research on flow in exercise.

Quantitative research on flow in exercise. In total, 25 studies used quantitative measures. The majority (*n* = 21) assessed flow intensity (i.e., the degree to which individuals experienced flow), while the remainder (*n* = 4) measured dispositional flow. Three measures were used to examine dispositional flow, but there was substantial heterogeneity in the measurement of flow intensity, with 16 measures used to assess this construct (see Table 2).

[INSERT TABLE 2 ABOUT HERE]

Psychometric properties of quantitative measures. In total, 68% (*n* = 17) of studies that used quantitative measures in the review used a validated psychometric inventory (i.e., included all items in the validated measure; see Table 2), while just under half (*n* = 11; 44%) reported information on the internal consistency of the measure used. The most frequently used validated measures were the: Flow State Scale (FSS; *n* = 3; Jackson & Marsh, 1996); Short Flow State Scale-2 (SFS; *n* = 3; Jackson, Martin, & Eklund, 2008); Flow Short Scale (*n* = 2; FShS; Rheinberg, Vollmeyer, & Engeser, 2003); Flow State Scale-2 (*n* = 2; Jackson & Eklund, 2004, 2002); and Dispositional Flow Scale-2 (DFS-2; *n* = 2; Jackson &

Eklund, 2004). As the majority of measures conceptualised flow based on the nine-dimensions framework (Csikszentmihalyi, 2002), this suggests that most knowledge stemming from quantitative research on flow in exercise is based on this framework.

Five studies (Grove & Lewis, 1996; Huang et al., 2018; Iida & Oguma, 2013; Lee, Myers, Park, Hill, & Feltz, 2018; Wollseiffen et al., 2016) removed items from validated measures, which raises concerns about the validity of the final measures employed. Additionally, four studies assessed flow using measures that were not previously validated. Two studies (Bronner, Pinsker, & Noah, 2015, 2013) combined items from a range of previously validated inventories, but did not assess the validity or reliability of the adapted measures. Kliem and Wiemeyer (2010) used items that “derived from” and “referred to” (p. 84) the game-flow model (Sweetser & Wyeth, 2005), but did not elaborate on the contents of the measure. Likewise, Park and Noh (2017) operationalised exercise flow “based on the theoretic background…of the experiment” (p. 1679), but did not provide information on the content or validity of the 5-item measure used.

Only one study in the review (Karageorghis et al., 2000) conducted a confirmatory factor analysis to test an a priori flow measurement model in exercise. Karageorghis et al. (2000) examined the factor structure of the FSS, which was previously validated in athletes (Jackson & Marsh, 1996). While Karageorghis et al. (2000) found that the nine-factor and hierarchical FSS measurement models demonstrated “satisfactory fit” and “reasonable fit” (p. 239), respectively, another study in exercisers reported that the unidimensional, nine-factor, and hierarchical FSS models displayed inadequate model fit (Vlachopoulos, Karageorghis, & Terry, 2000). Thus, caution should be taken when interpreting findings from studies that employed the FSS (Barry, van Schaik, MacSween, Dixon, & Martin, 2016; Karageorghis et al., 2000; Robinson, Dixon, MacSween, van Schaik, & Martin, 2015). Overall, just over two-thirds of included studies assessed flow using the full contents of validated measures and under half reported the internal consistency of the measurement tool used to assess flow. Collectively, only 24% of studies (*n* = 6) that employed quantitative measures used every item from a validated measure *and* examined the internal consistency of that measure in the study. Therefore, substantive questions exist regarding the psychometric properties of measures used in a large proportion of studies on flow in exercise.

Qualitative research on flow in exercise. The only qualitative study in the review (Swann et al., 2019) explored the experience, occurrence, and outcomes of flow in exercise through event-focused interviews (i.e., conducted within two days of the exercise activity on average). Swann et al. (2019) reported that flow was one of two[5] optimal psychological states experienced by exercisers (*n* = 18). Overall, findings reported by Swann et al. (2019) provide a multifaceted account of the flow experience in exercise, which encompassed the contexts, processes of occurrence, experience, and outcomes of flow (discussed below). While further work is required to substantiate and critically test this initial evidence, findings in relation to flow in the Integrated Model of Flow and Clutch States (Swann et al., 2019) appear to represent a promising attempt to explain the occurrence of flow in exercise.

Mixed method research on flow in exercise. Only one study adopted a mixed method design to investigate flow in exercise. Lee et al. (2016) collected quantitative data through the FSS-2 and qualitative data via verbal responses to open-ended questions on a questionnaire. However, as the quantitative data related to the exergame session as a whole and the qualitative data pertained to the eight individual exergames played during the session, the datasets were not reconcilable. Consequently, the quantitative and qualitative data were not integrated at any point in the analysis or interpretation of findings.

Conceptualisation of Flow in Exercise

[5] The second optimal psychological state experienced by exercise participants was defined as a 'clutch state' (Swann et al., 2019). Exercisers reported that flow and clutch states shared some features, but each state occurred through a different process, contained distinct elements, and produced unique outcomes (Swann et al., 2019).

A range of conceptualisations of flow have been employed in this field to date. These are discussed below in terms of two higher-order themes: (i) conceptual underpinning of quantitative measures; and (ii) conceptualisation of flow from qualitative research in exercise.

Conceptual underpinning of quantitative measures. Various conceptual frameworks have been applied to quantitative research on flow in exercise. The majority of studies (*n* = 17; 68%) that collected quantitative data used measures that conceptualised flow based on the nine-dimensions framework (Csikszentmihalyi, 2002). While Huang et al. (2018) adopted Hoffman and Novak's (1996) unidimensional model of flow, Hoffman and Novak (1996) stated that this conceptual model was influenced by the nine-dimensions framework (Csikszentmihalyi, 2002). Similarly, although Monedero et al. (2017) assessed "core flow", this construct is "consistent with original conceptualizations of subjective optimal experience underpinning flow" (Martin & Jackson, 2008, p. 150). In addition, several studies applied either the game-flow model (*n* = 3; Sweetser & Wyeth, 2005) or Rheinberg's (2008) flow model (*n* = 3), which both broadly reconcile with the nine-dimensions framework. Specifically, Rheinberg's (2008) conceptualisation of flow is an adaptation of the earliest description of flow (Csikszentmihalyi, 1975), with the autotelic experience omitted on the basis that this characteristic did not constitute a separate or additional component of flow. Likewise, the eight dimensions of the game-flow model, with the exception of social interaction, are proposed to "relate to" elements of Csikszentmihalyi's (2002) nine-dimensions framework (Sweetser & Wyeth, 2005, p. 3).

The two remaining quantitative studies did not conceptualise flow based on an existing framework. Park and Noh (2017) operationally defined "exercise flow" as a state of complete focus in which "unnecessary external information of thoughts do not enter the mind" (p. 1679), but did not explicate the conceptual underpinnings of the measure used. Grove and Lewis (1996) included flow within the 'flowlike states' construct based on overlapping characteristics between flow, peak experience, and peak performance. Therefore, given that flowlike states are conceptually broader than flow, this term will be used in reference to findings reported by Grove and Lewis (1996).

Conceptualisation of flow from qualitative research in exercise. Flow was characterised by 12 features in the qualitative study by Swann et al. (2019): absence of negative thoughts; absorption; altered perceptions; automatic skill execution; confidence; ease/reduced effort; effortless attention; enjoyment; feeling in control; optimal arousal; motivation for more; and positive-in-the-moment feedback. Although some of these characteristics broadly reconcile with those reported during flow previously, findings reported by Swann et al. (2019) represent an alternative conceptual model of flow in exercise. To date this is the only conceptual model of flow that has been developed through qualitative interviews with exercisers. Furthermore, this model also proposes that flow states share conceptual overlaps with another optimal "clutch" state, and that traditional conceptualisations of flow may not be able to distinguish between both (Swann et al., 2019).

Summary. The majority of quantitative studies on flow in exercise have employed measures based on the nine-dimensions framework (Csikszentmihalyi, 2002), which means that most understanding of flow in this setting is based on this conceptual model. However, given the lack of consensus regarding the conceptualisation of flow in studies that employed quantitative measures, this makes it difficult to amalgamate the findings. In addition, one qualitative study has reported an alternative conceptualisation of flow based on event-focused interviews with exercisers. While only a single study based on that conceptualisation has been conducted in exercise to date, this may be a promising avenue for future research.

Mechanisms of Occurrence and Strategies to Induce Flow in Exercise

A total of 18 studies examined mechanisms through which flow occurs or strategies to induce flow in exercise. The study designs ranged from experimental (*n* = 10), to longitudinal

($n = 4$), cross-sectional ($n = 3$), and qualitative ($n = 1$). Findings are discussed below in terms of six higher-order themes, comprising: (i) changes over time; (ii) virtual stimuli; (iii) exergame design features; (iv) music; (v) qualitative findings; and (vi) dispositional flow.

Changes over time. Seven studies obtained repeated measures of flow intensity over time, with six studies examining between-session differences and two studies assessing changes within exercise sessions. In longitudinal studies that obtained measures of global flow intensity in at least two exercise sessions, small increases from baseline to the final measurement point were found in Zumba ($d = 0.46$ - Elbe et al., 2016) and tai chi ($d = 0.21$ - Iida & Oguma, 2014) participants, while large increases were found in exergamers ($d = 0.95$ - Lee et al., 2018). Grove and Lewis (1996) found that the overall increase in circuit trainers from baseline was negligible ($d = 0.19$), although the length of time between measurement points (i.e., 6 weeks) was briefer than the aforementioned studies. Barry et al. (2016) observed no significant time-by-exercise interaction effects in exergamers and TGB exercisers for any of the nine FSS subscales in a 4-week intervention. However, small increases were evident in seven of the nine subscales in the exergamers ($0.22 \leq d\text{s} \leq 0.46$), although an effect was only found in two subscales for the TGB group ($d = 0.24$ and $d = 0.50$ - Barry et al., 2016). In a similar timeframe, Robinson et al. (2015) found moderate and large changes in all nine subscales from baseline in exergamers ($0.50 \leq d\text{s} \leq 1.24$), but small or moderate increases ($0.25 \leq d\text{s} \leq 0.62$) were only found in four subscales in the traditional balance training group.

Two studies examined within-session changes in global flow intensity during exercise. Grove and Lewis (1996) found a significant increase ($p < .0005$) in flowlike states in circuit trainers from the first to the second half of 45-50-minute sessions, although increases in flowlike states from the first half to the second half of sessions were greater and only significant ($d = 0.34$, $p < .0005$) in circuit trainers with high hypnotic susceptibility. In endurance running, flow increased from pre-run to the first hour (d = 0.79, $p < .05$) of a 6-hour run, before declining from the 1-hour to 3-hour (d = -0.54, $p < .05$) and 3-hour to 5-hour (d = -0.56) points (Wollseiffen et al., 2016).

Taken together, the findings from studies that assessed flow longitudinally highlight the dynamic nature of this psychological state. In terms of between-session changes, the findings offer cross-study evidence that continued engagement in exercise has a positive effect on flow intensity, or at least some dimensions of the flow experience. The findings from studies that examined within-session changes in flow intensity demonstrated the potential variability in flow intensity during exercise sessions. However, given the disparity in the length of the events examined, differences in findings, and potential influence of extraneous variables, this makes it difficult to draw strong inferences about the nature or extent of within-session changes in flow intensity at present.

Virtual stimuli. Five studies examined virtual stimuli and flow intensity by comparing levels of flow in exergaming to conventional exercise activities. Three studies used experimental approaches (Barry et al., 2016; Monedero et al., 2017; Robinson et al., 2015), while two employed cross-sectional designs (Lee et al., 2016; Thin, Hansen, & McEachen, 2011). A cross-over experimental study found a large effect of exercise type on core flow (η^2 = .52, $p < .05$), with higher scores reported in an entertainment-themed video game and a fitness-themed video game compared to treadmill running at moderate (d = 0.63 and d = 0.41, respectively) and self-selected intensities (d = 0.54 and d = 0.41, respectively; Monedero et al., 2017). In a randomised controlled trial that assessed flow intensity in the first and final session of 4-week TGB exercise and exergame interventions, Robinson et al. (2015) reported a main effect of group on all nine FSS subscales post-intervention, with higher scores reported in exergamers across eight of the nine subscales ($0.43 \leq d$s ≤ 1.71). Similarly, exergamers reported higher scores in all nine FSS subscales in the first ($0.44 \leq d$s $\leq$

0.96) and final ($0.66 \leq d$s ≤ 1.05) sessions of a 4-week intervention compared to a TGB exercise group (Barry et al., 2016).

Two cross-sectional studies examined differences in flow intensity between exergamers and TGB exercise participants. Both studies (Lee et al., 2016; Thin et al., 2011) compared scores on the FSS-2 in small samples of exergamers (n = 25 and n = 14, respectively) to previously published normative scores in exercisers (n = 200 - Jackson & Eklund, 2004) but reported different findings. Although each study reported significantly higher ($p < .05$) scores for two of the nine FSS-2 subscales in exergamers, there was no consistency in terms of the dimensions that displayed significant differences.

Taken together, evidence from higher quality studies suggests that virtual stimuli could be associated with higher values in at least some dimensions of flow. However, given that 80% of studies that investigated differences between traditional exercise and exergaming failed to analyse flow at a global level (i.e., obtaining a total score for flow), it is difficult to make firm conclusions about the relationship between virtual stimuli and global flow intensity.

Exergame design features. Five experimental studies examined the effect of exergame design features on flow intensity, with all five studies employing a different measure of flow. Compared to a control condition (i.e., cycling exergame without a software generated partner [SGP]), exergaming with a SGP that was not always superior had a large and significant positive effect on flow at both the second (d = 1.12, p = .029) and third (d = 1.17, p = .021) measurement points in a 24-week exergame intervention, whereas exergaming with an SGP that was always superior did not have an effect at the second (d = 0.62, p = .182) or third (d = 0.01, p = .985) blocks (Lee et al., 2018). Bronner et al. (2015, 2013) reported a significant difference in game-flow between exergames (η_p^2 = .49, p = .002 and p = .025, respectively). Post hoc comparisons were non-significant in each sample (both n = 7), but small, moderate, and large differences were evident between exergames (0.22 ≤ *d*s ≤ 2.12 - Bronner et al., 2015, 2013). Park and Noh (2017) reported higher exercise flow ($d = 0.53$, $p < .01$) in an "enhanced interactivity" condition, which included "exercise command narration", compared to a "weaker interactivity" condition (p. 1679). However, little information was provided on the content of the 5-item flow measure, or the nature of the 10-minute exercise activities, which raises concerns about the results. Finally, higher core flow ($d = 0.26$) was reported in an entertainment-themed exergame compared to a fitness-themed exergame, although the difference was not significant (Monedero et al., 2017). Overall, due to the limited data available, the degree of variation in design features examined, and heterogeneity in measurement approaches, it is difficult to reconcile the findings. Nevertheless, there is cross-study evidence that the design of an exergame affects flow intensity.

Music. Two experimental studies examined the effect of music on global flow intensity during treadmill exercise. Karageorghis et al. (2008) found a large, significant main effect ($\eta_p^2 = .49$, $p = .000$) of music tempi on flow intensity, with participants reporting significantly higher flow intensity ($p = .000$) in fast tempi ($d = 1.26$), medium tempi ($d = 1.40$), and mixed tempi ($d = 1.64$) music conditions compared to a no-music control condition. Similarly, Karageorghis and Jones (2014) reported significant differences in flow across music tempo conditions ($\eta_p^2 = .28$, $p = .002$) in treadmill walking/running, with flow intensity significantly higher ($p < .05$) in the slow ($d = 0.46$), medium ($d = 0.76$), fast ($d = 0.79$), and very fast ($d = 0.50$) tempi conditions compared to the no-music condition. While it is difficult to draw firm conclusions about knowledge based on two studies, the available evidence suggests that music has a positive effect on flow intensity in walking/running.

Qualitative findings. Two studies reported qualitative data on the occurrence of flow in exercise. In the only mixed method study in the review, Lee et al. (2016) collected qualitative data by asking participants to identify which of eight 2.5-minute exergames produced the “most” and “least” (p. 244) flow and then to verbally explain the reason underlying these selections. The most common reasons for experiencing the most flow in specific exergames were ‘correct level of difficulty’ and ‘clear goals’, while ‘inadequate level of difficulty’ and ‘pain’ were reported most frequently as reasons for experiencing flow the least (Lee et al., 2016). However, no information was provided regarding how these factors influenced the level of the flow experience, thus providing little insight into the potential causal mechanisms underpinning the occurrence of flow states. The only qualitative study in the review reported that flow states occurred in exercise contexts that involved flexible outcomes, exploration, and novelty/variation (Swann et al., 2019). In such contexts, the process of flow occurrence was initiated by a positive event, which provided positive feedback. This feedback increased confidence and led exercisers to challenge themselves and set open goals (i.e., that did not include specific outcomes), which enabled the transition into flow. These findings suggest that flow occurs through a combination of mechanisms, which, together, may constitute necessary and sufficient conditions of flow. While only one study has been conducted from this perspective in exercise, the findings may provide useful avenues for future research seeking to develop an explanation of flow.

Dispositional flow. Two studies examined effects of intervention programmes on dispositional flow. Kliem and Wiemeyer (2010) found no significant difference in game-flow between participants in 3-week balance training and exergaming programmes. However, as this study employed a measure of game-flow for both samples, it is arguable that this might not have been as appropriate for the traditional balance training group given that this construct stems from the computer-gaming literature. In turn, this raises concerns with the quality of this evidence. Butzer et al. (2016) examined changes in dispositional flow at the beginning and end of an 8-week yoga intervention in adult musicians and found a small improvement in dispositional flow over time (d = 0.48). With evidence only drawn from two studies, this illustrates that there is limited knowledge of mechanisms of occurrence or strategies for enhancing dispositional flow in exercise.

Correlates and Outcomes of Flow in Exercise

A total of 15 studies reported correlates and factors associated with flow, while five studies examined outcomes related to flow. These findings are discussed below in terms of two higher-order themes: (i) correlates of flow in exercise, within which there are a number of lower-order themes; and (ii) outcomes associated with flow in exercise.

Correlates of flow in exercise. The findings relating to correlates of flow are presented in terms of six lower-order themes: (i) physiological correlates of flow states; (ii) gender; (iii) age; (iv) exercise experience; (v) additional findings pertaining to flow intensity; and (vi) correlates of dispositional flow.

Physiological correlates of flow states. Five studies reported data on flow intensity and physiological indices. Grove and Lewis (1996) found a non-significant relationship between self-reported heart rate and flowlike states in circuit training. Bronner et al. (2013) found a positive association between game-flow and VO_2 ($r = .57$), while two studies (Bronner et al., 2015, 2013) found a moderate association ($r = .52$ and $r = .59$, respectively) between metabolic equivalent of the task (MET) and game-flow in small exergaming samples ($n = 7$). In contrast, while the relationship between MET and flow was not statistically tested, Monedero et al. (2017) reported that core flow was highest in the condition with the lowest MET output, which equated to moderate-intensity exercise. In previously inactive males, no significant relationship was found between global flow intensity in a mid-intervention session and future physiological improvements after 12-week strength or interval running interventions (Elbe et al., 2010). Collectively, findings pertaining to the physiological correlates of flow in exercise are equivocal. Furthermore, due to the variance in the physiological and flow measures employed, it is difficult to draw firm conclusions at present.

Gender. Three studies investigated differences in flow intensity between men and women. Wollseiffen et al. (2016) found that women (n = 5) reported higher flow intensity than men (n = 6) before and during a 6-hour run ($0.80 \leq d$s ≤ 2.00). In contrast, a cross-sectional study only found significantly higher scores (d = 0.64, p = .028) for women in two of the 26 Activity Flow State Scale (AFSS) items in exergaming (Marston et al., 2016), thus suggesting negligible differences. While men reported higher flow intensity in a no-music control condition (d = 1.04) in treadmill walking/running, negligible differences were found in the slow and very fast music tempi conditions, while only small differences were found in the medium and fast conditions, with women reporting higher flow intensity (d = 0.26 and d = 0.27, respectively - Karageorghis & Jones, 2014). In short, based on the small sample size (Wollseiffen et al., 2016) and equivocal nature of the findings (Karageorghis & Jones, 2014; Marston et al., 2016) it appears that gender differences in flow intensity are negligible.

Age. Two cross-sectional studies examined the relationship between flow intensity and age in exergaming. Lee et al. (2016) found no significant relationship between flow intensity and age in a single exergame session. Similarly, Marston et al. (2016) observed no significant difference in flow intensity between different age groups in exergaming. Although the evidence stems from two small samples (n = 25 and n = 50, respectively), there is tentative evidence that age is not significantly related to flow intensity in exergaming.

Exercise experience. Two studies assessed the relationship between exercise experience and flow intensity. Grove and Lewis (1996) found that experienced circuit trainers (i.e., ≥ six months experience) reported higher flowlike states (d = 0.43, p < .021) compared to their less experienced counterparts. A cross-sectional study found a significant positive relationship (β = .33, p < .001) between years of tai chi experience and global flow intensity (Iida & Oguma, 2013). While based on a small number of studies, the findings offer tentative evidence that prior experience is positively associated with flow intensity.

Additional findings pertaining to flow intensity. Several other findings in relation to flow intensity were examined in single studies and were subsequently not coded into themes. A cross-sectional study found no significant relationship between flow intensity and either physical dysfunction or perceptions of knee pain in exergamers (Lee et al., 2016). Marston et al. (2016) reported significant differences in 19 of the 26 AFSS items between three international study sites, but found no significant differences in global flow intensity or the nine dimensions of flow between locations. On this basis, caution should be taken when considering the significant findings reported. Another cross-sectional study found a significant, but small, positive relationship between flow intensity and perceived challenge ($r = 0.21$, $p < .05$), but no significant relationship was observed between flow intensity and either need for exercise or need for achievement (Huang et al., 2018). Finally, an online, questionnaire study in exergamers reported that flow intensity was significantly predicted by game level ($\beta = .25$, $p < .001$), playing alone ($\beta = .23$, $p = .002$), nostalgia ($\beta = .20$, $p = .002$), and playing with family ($\beta = .15$, $p = .024$), but was not significantly predicted by playing with friends or strangers (Loveday & Burgess, 2017). Overall, as each of these findings are drawn from single studies, it is difficult to generate firm conclusions about them at present.

Correlates of dispositional flow. Two quantitative studies used cross-sectional designs to examine factors associated with dispositional flow. Ersöz et al. (2016) found significant differences ($p < .05$) in global dispositional flow between underweight, normal, and overweight and obese exercisers, with higher scores observed in underweight participants compared to overweight and obese participants. Ersöz and Eklund (2017) found no significant differences ($-0.20 \leq d$s ≤ 0.17) in any of the nine DFS-2 subscales between genders but reported a significant multivariate effect for stage of change in exercise on the DFS-2 subscales ($\eta_p^2 = .075$, $p < .05$). However, as significant effects were only apparent in three of the nine subscales ($.03 \leq \eta_p^2 \leq .08$, $p \leq .02$), this provides little evidence of a relationship between stages of changes and global dispositional flow. Further, Ersöz and Eklund (2017) found that some dimensions of dispositional flow were significantly ($p < .05$) and positively related to intrinsic regulation, identified regulation, and introjected regulation, and inversely associated with external regulation, and amotivation. However, only intrinsic regulation was significantly related to all nine dimensions ($.27 \leq r\text{s} \leq .66$), while only one dimension of flow, autotelic experience, displayed a significant association ($p < .05$) with intrinsic regulation ($r = .66$), identified regulation ($r = .46$), introjected regulation ($r = .26$), external regulation ($r = -.19$), and amotivation ($r = -.27$). Collectively, as there is only a single study to draw upon for each finding in relation to correlates of dispositional flow, this demonstrates a lack of evidence in this area. In turn, this makes it difficult to generate meaningful conclusions about correlates of dispositional flow in exercise at present.

Outcomes associated with flow in exercise. Five studies quantitatively examined relationships between flow intensity and various outcome variables. In a longitudinal study, Iida and Oguma (2014) reported that global flow intensity in a single tai chi class significantly predicted ($\beta = .18$, $p = .019$) sense of coherence one year later in experts, but not in non-experts or the entire sample. A cross-sectional study in aerobic dance exercisers (Karageorghis et al., 2000) found significant, moderate-to-strong, positive associations between flow intensity and post-exercise feelings of positive engagement ($\beta = .59$, $p < .05$), revitalisation ($\beta = .55$, $p < .05$), and tranquillity ($\beta = .46$, $p < .05$), and a significant, but weak, negative relationship with physical exhaustion ($\beta = -.12$, $p < .05$). In addition, flow intensity explained significant variance ($p < .01$) in positive engagement, revitalisation, and tranquillity (35%, 31%, and 22%, respectively), but accounted for little variance in physical exhaustion (Karageorghis et al., 2000). In a cross-sectional study, Iida and Oguma (2013) found that flow intensity in tai chi was significantly and positively associated ($\beta = .51$, $p < .05$) with 'ikigai', which is regarded as an indicator of wellbeing in Japanese culture (e.g., Sone et al., 2008). Another cross-sectional study found a significant, but small, positive relationship ($r = 0.18$, $p < .05$) between flow and exercise enjoyment (Huang et al., 2018). In a qualitative study (Swann et al., 2019), exercisers reported a range of positive consequences after the activity, including: a sense of achievement; confidence; energy; intrinsic motivation; and positive mood and emotions. Taken together, there is tentative evidence that flow is associated with a range of positive and desirable outcomes. However, given that these findings primarily stem from cross-sectional studies, it is not possible to infer causality.

Discussion

The purpose of this study was to systematically review, synthesise, and appraise current knowledge on flow in exercise settings. Research on flow in exercise has only been conducted relatively recently, with the first study published in 1996 (Grove & Lewis, 1996). Since then, the area has grown and an in-depth body of research has developed. Indeed, this study illustrates that research on flow in exercise has grown rapidly in recent years, with the vast majority ($n = 23$; 88%) of included studies published in or after 2010. In contributing the first systematic review, synthesis, and appraisal of flow research in exercise settings, this study aimed to address the following questions: (i) how has flow been measured in exercise?; (ii) how has flow been conceptualised in exercise?; (iii) what is the quality of research in this field?; (iv) what evidence is available regarding mechanisms of occurrence and strategies for inducing flow?; and (v) what are the correlates of flow in exercise?

Summary of Findings

The majority of studies in the review collected data on flow using quantitative measures, with only two studies (8%) obtaining qualitative data. Most quantitative studies in the review conceptualised flow based on the nine-dimensions framework (Csikszentmihalyi, 2002) or conceptually similar models (Rheinberg, 2008; Sweetser & Wyeth, 2005). Findings concerning the experience of flow in exercise generated by the only qualitative study in the review (Swann et al., 2019) largely reconcile with these frameworks, but do offer a slightly different perspective on the conceptualisation of flow in this context. Overall, quantitative measures and the nine-dimensions framework (Csikszentmihalyi, 2002) have been widely used to study flow in exercise, but the findings of this review highlight a range of conceptual and methodological issues in the field (see below). In turn, this raises some doubts about the quality of this evidence base.

Flow has been studied across an array of exercise activities and the findings of this review demonstrate that a broad evidence base is apparent. However, there was limited data available in relation to a range of themes identified in the review, which makes it difficult to draw firm conclusions about many aspects of this literature at present. Nevertheless, there was some tentative, cross-study evidence (i.e., similar findings in ≥ 2 studies) regarding several findings, particularly in relation to flow intensity. In terms of mechanisms of occurrence and strategies for inducing flow, there was cross-study evidence from a number of higher-quality, experimental studies that flow intensity, or at least some dimensions of this phenomenon, could be influenced by: the presence of virtual stimuli during exercise (i.e., in exergames; Barry et al., 2016; Monedero et al., 2017; Robinson et al., 2015); design features of exergames (Bronner et al., 2015, 2013; Lee et al., 2018; Monedero et al., 2017; Park & Noh, 2017); and music (Karageorghis & Jones, 2014; Karageorghis et al., 2008). Furthermore, although based on limited data, recent findings from qualitative research (Swann et al., 2019) offer potential insights into the contexts and process underlying the occurrence of flow in exercise, which could provide an avenue to move towards developing an explanation for flow in this context.

In observational studies, some tentative evidence, albeit from a limited number of studies, suggests that exercise experience (e.g., years participating in exercise) is positively related to flow intensity (Grove & Lewis, 1996; Iida & Oguma, 2013). A noteworthy finding from this review was that flow has been associated with a variety of positive outcomes in exercisers, including: a sense of achievement; confidence; positive energy and emotions after exercise; and intrinsic motivation. Given the desirability and relevance of these outcomes for exercisers of all ages and standards, the findings of the review support the contention that flow could offer a potential avenue to promote long-term engagement in physical activity (Petosa & Holtz, 2013). While further work is required to determine whether flow is a causal determinant of exercise adherence, findings concerning the potential behavioural and psychological benefits of flow states in exercise highlight the importance of developing the exercise-specific knowledge base on this experience to enable exercisers to achieve this state and its associated outcomes more reliably and consistently.

Conceptual Issues

Findings of this review reveal a number of conceptual issues in flow research in exercise. First, there is considerable variation in the conceptualisation of flow applied to quantitative research in exercise. While the majority of studies adopted the nine-dimensions framework (Csikszentmihalyi, 2002), two additional conceptual models (Rheinberg, 2008; Sweetser &Wyeth, 2005) were also employed. Furthermore, Grove and Lewis (1996) incorporated flow within another construct, ‘flowlike states’, based on overlapping characteristics between flow, peak experience, and peak performance in sport, which is problematic given that these three phenomena are regarded as distinct optimal experiences (Jackson & Kimiecik, 2008). Although the various models concerning flow research in exercise share some common conceptual ground, several differences are apparent (e.g., in the dimensions of flow), which makes it difficult to interpret and reconcile findings in this area.

Second, flow is most commonly conceptualised based on the nine-dimensions framework (Csikszentmihalyi, 2002), but this conceptualisation has been subjected to extensive criticism (e.g., Drengner, Jahn, & Furchheim, 2018; Hassmén, Keegan, & Piggott, 2016; Jackman, Fitzpatrick, Lane, & Swann, 2019). Within sport and exercise, concerns have been raised about the lack of precision in flow dimensions, overlapping and missing dimensions, low endorsement of some characteristics, and uncertainty regarding the number of dimensions that must be present to constitute a flow state (see Swann et al., 2018). Due to the ambiguity surrounding Csikszentmihalyi's (2002) framework, Swann et al. (2018) argued that "it is difficult to confidently proceed with the traditional nine-dimension paradigm" (p. 259). Thus, there are doubts about the most common conceptualisation of flow used in exercise research. Against this backdrop of uncertainty, there is a clear need for researchers to question fundamental assumptions concerning the conceptualisation of flow in exercise and to move towards developing a more refined conceptualisation of the phenomenon in this context.

Third, a noteworthy finding in this review was that no measure developed or validated specifically in exercise[6] was employed to assess flow. Prior to proceeding with further research on a new phenomenon, it is advisable to conduct qualitative research to check whether existing theoretical frameworks suitably describe the phenomenon (Giner-Sorolla, 2019). However, the first and only qualitative study to explore what it is like to experience flow in exercise (Swann et al., 2019) was published *after* each of the remaining 25 studies included in this review. While more than half of studies that employed quantitative measures evaluated measurement qualities from a statistical perspective (i.e., by examining internal consistency and/or conducting confirmatory factor analysis), little consideration has been given to critically testing the quality of the measurement tools used in exercise from a conceptual perspective. Indeed, of those that did critically test an a priori measurement model of flow in exercise, one raised concerns (Vlachopoulos et al., 2000) and the other only

[6] Although samples involved in the validation of the FSS-2 and DFS-2 included exercise participants (Jackson & Eklund, 2002), their data were combined with dance and sport participants for data analysis.

offered tentative support for the nine-dimensions framework (Karageorghis et al., 2000). Taken together, the evidence suggests that there are some questions about the validity of existing measures used to assess flow in exercise. Therefore, new or refined measures are arguably required to re-examine and extend knowledge of flow in exercise.

Theoretical Explanation for Flow States

An important finding from this review is that little research has been directed towards developing a causal explanation for the occurrence of flow in exercise. As pointed out by Kimiecik and Stein (1992) "it is one thing to know, for example, that a flow experience is accompanied by focused concentration, feelings of control, and clear goals. It is quite another to know why or how the flow experience actually occurred…The former emphasizes description; the latter focuses on the mechanisms underlying the experience" (p. 148). Similar to research in sport (see Swann, Piggott, Crust, Keegan, & Hemmings, 2015), most understanding of flow occurrence in exercise is based on *association* rather than *explanation*. Only one study in this review explicitly sought to investigate the causal mechanisms underlying the occurrence of flow in exercise (Swann et al., 2019). Although numerous quantitative studies investigated flow using experimental designs, flow was typically assessed as a secondary outcome (i.e., of an independent variable) and none of these studies explicitly sought to induce flow based on empirical findings or an explanatory theory for flow.

While some experimental studies found positive effects of different variables on flow intensity (i.e., with flow as a secondary outcome), few attempts were made to explain *how* and *why* these variables might have influenced or could be involved in the occurrence of flow. Occasionally researchers offered tentative proposals, with some suggesting that the immersive nature (Barry et al., 2016) and distractive elements (e.g., Robinson et al., 2015) of the activity could facilitate specific dimensions of flow. In contrast, Elbe et al. (2010) suggested that "the heart rates of the participants were kept rather high, which might be an explanation for the high experiences of flow" (p. 116), although the relationship between heart rate and flow was not tested and no information on the mechanisms underlying this proposed explanation was advanced. Some researchers alluded to the challenge-skills balance proposition (see Fong, Zaleski, & Leach, 2015 for a meta-analysis) by suggesting that flow was facilitated by perceiving the difficulty of a task as being a manageable challenge (Lee et al., 2018) or equal to personal skills (Lee et al., 2016). However, while challenge-skills balance might be an important factor underlying the occurrence of flow, it is argued that this condition alone is not sufficient for flow states to occur (Swann et al., 2018). Taken together, the findings suggest that further research that seeks to develop an understanding of the causal mechanisms underlying the occurrence of flow in exercise is required.

Methodological Critique

Developing valid and reliable methods to assess subjective experiences is one of the greatest challenges for flow researchers (Jackson & Kimiecik, 2008) and a number of methodological issues are apparent in flow research in exercise. First, although quantitative and qualitative methods can each provide unique understanding of flow (Jackson & Kimiecik, 2008), 96% (*n* = 25) of included studies employed quantitative measures while only 8% (*n* = 2) collected qualitative data (i.e., one study used quantitative and qualitative methods). Further, it should be noted that qualitative data collected by Lee et al. (2016) as part of a mixed method study were obtained through open-ended survey questions, which may have limited the depth of information acquired. Indeed, given that interviews can yield rich, detailed insights into lived experiences (e.g., Jackson & Kimiecik, 2008; Swann et al., 2012), it is surprising that only one interview study (Swann et al., 2019) has been conducted on flow in exercise in over two decades of research. Thus, given the limited amount of qualitative data collected to date, it is possible that valuable insights have yet to be attained.

Second, there is considerable heterogeneity in the quantitative measurement of flow in exercise, which makes it difficult to reconcile findings in this area. Eighteen quantitative measures were used in 25 studies, with this variety highlighted most ostensibly in exergaming research, where 12 quantitative measures were used across 13 studies. The review raises substantive questions with the psychometric properties of measures employed to study flow in many studies, with some of the most noteworthy issues including the: use of measures that were previously found to display inadequate model fit; development of measures to assess flow without testing their validity; and absence of information about the items used to represent flow. Further, the measurement tools employed by flow researchers in exercise are underpinned by a range of conceptual frameworks (e.g., Csikszentmihalyi, 2002; Rheinberg, 2008; Sweetser & Wyeth, 2005), which means that literature in this area was not only generated through different measures but is also based on different conceptualisations of flow. In turn, this diversity creates a somewhat confusing picture and makes it difficult to form a clear judgement about findings in this area to date.

Finally, it is interesting to note the relative absence of qualitative research on flow in exercise compared to sport, where this methodology has been used in many classic studies (e.g., Jackson, 1995, 1996). Notably, recent findings from the only qualitative study on flow in this review (Swann et al., 2019) offer an alternative perspective on what it is like to experience flow during exercise compared to the conceptual models underpinning the majority of quantitative research on flow in this domain. That description of flow in exercise was consistent with studies in sport (Jackman, Crust, & Swann, 2017; Swann et al., 2017a, 2017b; Swann, Crust, & Vella, 2017; Swann, Keegan, Crust, & Piggott, 2016), and suggests that a “clutch” state is conceptually similar to flow, but occurs in different contexts and through other processes (Swann et al., 2019). In turn, this raises further questions about the discriminant validity of quantitative measures used in flow research in exercise to date

(Swann et al., 2018). Further, it is important to note that Swann et al. (2019) employed 'event-focused' interviews, which were conducted within two days of the participant's experience on average, thus enabling a more detailed examination of the contexts, process of occurrence, experience, and outcomes of flow states in exercise. As such, this method may be a useful avenue for future research.

Limitations

This study attempted to systematically review a growing literature on flow in exercise. Although there are several strengths in the current review, including the strict definition of eligible participants and employment of trustworthiness procedures (e.g., critical friends, article screening, and assessment of study quality by multiple authors), it is important to note several issues with the approach employed. Some findings in this review are based on as little as one or two papers, which makes it difficult to draw firm conclusions about the evidence base. The eligibility criteria might have excluded participants that could have been relevant to the aim of this review but did not satisfy all of the inclusion/exclusion criteria. For example, this review only included peer-reviewed articles in the English language and might therefore be affected by a language bias and publication bias. Similarly, the review excluded potentially relevant participants due to the merging of sport and exercise participants in several studies. Further, it is also possible that relevant studies could have been missed in the electronic database search due to the search terms that were employed (e.g., by including the term "psycholog*"). While the limitations of this review are acknowledged, some of these issues (e.g., basing findings on a limited number of studies) could not be circumvented due to the nature of research pertinent to this review.

Future Research Directions

Based on the findings of this review, a series of recommendations are proposed to improve future research and advance understanding of flow in exercise. First, there is a need to assess the validity of flow measures used in exercise to: (i) enable a more informed evaluation of current knowledge on flow in exercise; (ii) establish whether existing measures represent the flow experience in exercise (i.e., assess construct validity and discriminant validity) and should continue to be used; and (iii) identify if there is a need to develop new measures for flow in this context. Second, future studies should continue to obtain real-time measures of flow during exercise (e.g., Grove & Lewis, 1996) to capture the complex and dynamic nature of this state and permit more fine-grained analysis. While it has been acknowledged that it is impractical to interrupt an athlete during sport to obtain such data (Jackson & Kimiecik, 2008), exercise activities, which often include scheduled breaks (e.g., between sets), could offer a more practical setting in which to collect real-time data (e.g., self-report and physiological) on flow. Indeed, by obtaining in-the-moment data on flow states during exercise activities, this could help to advance understanding about the conceptualisation and occurrence of flow in exercisers. Third, in future, researchers using componential measures should analyse flow at both a global and subscale level, and ensure that findings at a subscale level are not interpreted as being representative of the flow experience. Fourth, researchers should seek to understand whether questionnaire scores can be used to identify exercisers who did or did not experience flow and determine the most valid and reliable method to do so. Finally, and most importantly, further research should be explicitly directed towards understanding the causal mechanisms underpinning the occurrence of flow states in exercise, and thus move towards developing an explanatory theory for this phenomenon (cf. Swann et al., 2018). At present, the lack of explanation and understanding of causal mechanisms underlying the occurrence of flow makes it difficult to develop experiments or design evidence-based interventions that can reliably induce flow. By taking steps towards developing an explanatory theory, this could provide a more robust platform to develop practical recommendations that allow exercise participants to experience flow and its associated positive outcomes more reliably and consistently.

References

Notes: (1) Author details and references for validated quantitative measures used by included studies are available in Supplementary File 5.

Barry, G., van Schaik, P., MacSween, A., Dixon, J., & Martin, D. (2016). Exergaming (XBOX Kinect™) versus traditional gym-based exercise for postural control, flow and technology acceptance in healthy adults: a randomised controlled trial. *BMC Sports Science, Medicine and Rehabilitation, 8*(25), 1-11. doi: 10.1186/s13102-016-0050-0

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2015). *Comprehensive Meta-Analysis: Version 3.0 [Computer Program]*. Englewood, NJ: Biostat.

Bronner, S., Pinsker, R., & Noah, J. A. (2013). Energy cost and game flow of 5 exer-games in trained players. *American Journal of Health Behaviour, 37*(3), 369-380. doi: 10.5993/AJHB.37.3.10

Bronner, S., Pinsker, R., & Noah, J. A. (2015). Physiological and psychophysiological responses in experienced players while playing different dance exer-games. *Computers in Human Behaviour*, *51*, 34-41. doi: 10.1016/j.chb.2015.04.047

Butzer, B., Ahmed, K., & Khalsa, S. B. S. (2016). Yoga enhances positive psychological states in musicians. *Applied Psychophysiology and Biofeedback, 41*, 191-202. doi: 10.1007/s10484-015-9321-x

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126-131.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety.* San Francisco, CA: Jossey-Bass, Inc.

Csikszentmihalyi, M. (2002). *Flow: The classic work on how to achieve happiness*. London: Rider Books.

Drengner, J., Jahn, S., & Furchheim, P. (2018). Flow revisited: process conceptualisation and a novel application to service contexts. *Journal of Service Management*, *29*(4), 703-734. doi: 10.1108/JOSM-12-2016-0318

Elbe, A.-M., Barene, S., Strahler, K., Krustrup, P., & Holtermann, A. (2016). Experiencing flow in a workplace physical activity intervention for female health care workers: A longitudinal comparison between football and Zumba. *Women in Sport and Physical Activity, 24*, 70-77. doi: 10.1123/wspaj.2015-0011

Elbe, A.-M., Strahler, K., Krustrup, P., Wikman, J., & Stelter, R. (2010). Experiencing flow in different types of physical activity intervention programs: three randomized studies. *Scandinavian Journal of Medicine and Science in Sports, 20*(Suppl. 1), 111-117. doi: 10.1111/j.1600-0838.2010.01112.x

Ersöz, G., & Eklund, R. C. (2017). Behavioural regulations and dispositional flow in exercise among American students relative to stages of change and gender. *Journal of American College Health, 65*(2), 94-102. doi: 10.1080/07448481.2016.1239203

Ersöz, G., Altiparmak, E., & Hülya Aşçı, F. (2016). Does body mass index influence behavioural regulations, dispositional flow and social physique anxiety in exercise setting? *Journal of Sports Science and Medicine, 15*, 295-300.

Fong, C. J., Zaleski, D. J., & Leach, J. K. (2015). The challenge–skill balance and antecedents of flow: A meta-analytic investigation. *The Journal of Positive Psychology, 10*(5), 425-446.

Giner-Sorolla, R. (2019). The past thirty years of emotion research: appraisal and beyond. *Cognition and Emotion, 33*(1), 48-54. doi: 10.1080/02699931.2018.1523138

Grove, J. R., & Lewis, M. A. E. (1996). Hypnotic susceptibility and the attainment of flowlike states during exercise. *Journal of Sport & Exercise Psychology*, *18*, 380-391. doi: 10.1123/jsep.18.4.380

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature, 555,* 175-182. 10.1038/nature25753.

Hassmén, P., Keegan, R., & Piggott, D. (2016). *Rethinking sport and exercise psychology research*. London, UK: Springer Nature.

Hoffman, D. L., & Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing, 60,* 50-68. doi: 10.2307/1251841

Huang, H-C., Pham, T. T. L., Wong, M-K., Chiu, H-Y., Yang, Y-H., & Teng, C-I. (2018). How to create flow experience in exergames? Perspective of flow theory. *Telematics and Informatics, 35,* 1288-1296. doi: 10.1016/j.tele.2018.03.001

Iida, K., & Oguma, Y. (2013). Relationships between flow experience, IKIGAI and sense of coherence in tai chi practitioners. *Holistic Nursing Practice, 27*(5), 260-267. doi: 10.1097/HNP.0b013e31829b9199

Iida, K., & Oguma, Y. (2014). The relationship between flow experience and sense of coherence: A 1-year follow-up study. *Holistic Nursing Practice, 28*(2), 91-97. doi: 10.1097/HNP.0000000000000018

Jackman, P. C., Crust, L., & Swann, C. (2017). Systematically comparing methods used to study flow in sport: A longitudinal multiple-case study. *Psychology of Sport & Exercise, 32*, 113–123. doi: 10.1016/j.psychsport.2017.06.009

Jackman, P. C., Fitzpatrick, G., Lane, A., & Swann, C. (2019). Exploring bodily sensations experienced during flow states in professional national hunt jockeys: a connecting analysis. *Qualitative Research in Sport, Exercise and Health, 11*(1), 92-105. doi: 10.1080/2159676X.2017.1380693

Jackson, S. (1995). Factors influencing the occurrence of flow state in elite athletes. *Journal of Applied Sport Psychology, 7*, 138–166. doi: 10.1080/10413209508406962

Jackson, S., & Eklund, R. (2002). Assessing flow in physical activity: The Flow State Scale-2 and Dispositional Flow Scale-2. *Journal of Sport & Exercise Psychology*, *24*, 133-150.

Jackson, S., & Eklund, R. (2004). *The flow scales manual*. Morganstown, WV: Fitness Information Technology.

Jackson, S., & Kimiecik, J. (2008). Optimal experience in sport and exercise. In T. Horn (Ed.), *Advances in sport psychology* (pp. 377- 399). Champaign, IL: Human Kinetics.

Jackson, S., & Marsh, H. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of Sport & Exercise Psychology*, *18*, 17-35. doi: 10.1123/jsep.18.1.17

Jackson, S., (1992). Athletes in flow: A qualitative investigation of flow states in elite figure skaters. *Journal of Applied Sport Psychology, 4,* 161-180. doi: 10.1080/10413209208406459

Jackson, S., Martin, A. J., & Eklund, R. C. (2008). Long and short measures of flow: Examining construct validity of the FSS-2, DFS-2 and new brief counterparts. *Journal of Sport & Exercise Psychology*, *30*, 561-587.

Karageorghis, C. I., & Jones, L. (2014). On the stability and relevance of the exercise heart rate-music-tempo preference relationship. *Psychology of Sport and Exercise, 15,* 299-310. doi: 10.1016/j.psychsport.2013.08.004

Karageorghis, C. I., Vlachopoulos, S. P., & Terry, P. C. (2000). Latent variable modelling of the relationship between flow and exercise-induced feelings: An intuitive appraisal perspective. *European Physical Education Review, 6*(3), 230-248. doi: 10.1177/1356336X000063002

Karageorghis, C., Jones, L., & Stuart, D. P. (2008). Psychological effects of music tempi during exercise. *International Journal of Sports Medicine, 29*(7), 613-619. doi: 10.1055/s-2007-989266

Kimiecik, J. C., & Harris, A. T. (1996). What is enjoyment? A conceptual/definitional analysis with implications for sport and exercise psychology. *Journal of Sport & Exercise Psychology, 18,* 247-263. doi: 10.1123/jsep.18.3.247

Kimiecik, J., & Stein, G. (1992). Examining flow experiences in sport contexts: Conceptual issues and methodological concerns. *Journal of Applied Sport Psychology*, *4*, 144–160. doi: 10.1080/10413209208406458

Kliem A. & Wiemeyer, J. (2010). Comparison of a traditional and a video game-based balance training program. *International Journal of Computer Science in Sport*, *9*(Special Edition), 80-91.

Lee, M., Suh, D., Son, J., Kim, J., Eun, S-D., & Yoon, B. (2016). Patient perspectives on virtual reality-based rehabilitation after knee surgery: Importance of level of difficulty. *Journal of Rehabilitation Research and Development, 53*(2), 239-252. doi: 10.1682/JRRD.2014.07.0164

Lee, S., Myers, N. D., Park, T., Hill, C. R., & Feltz, D. L. (2018). An exploratory study on the Köhler effect and flow in long-term exergaming. *Simulation and gaming*, *49*(5), 538-552. doi: 10.1177/1046878118776043

Loveday, P., & Burgess, J. (2017). Flow and Pokémon GO: The contribution of game level, playing alone, and nostalgia to the flow state. *E-Journal of Social & Behavioural Research in Business*, *8*(2), 16-28.

Maddison, R., Mhurchu, C. N., Jull, A., Jiang, Y., Prapavessis, H., & Rodgers, A. (2007). Energy expended playing video console games: an opportunity to increase children's physical activity? *Pediatric exercise science*, *19*(3), 334-343. doi: 10.1123/pes.19.3.334

Marston, H. R., Kroll, M., Fink, D., & Gschwind, Y. J. (2016). Flow experience of older adults using the iStoppFalls exergame. *Games and Culture, 11*(1-2), 201-220. doi: 10.1177/1555412015605219

Milner, K. A. (2015). Systematic reviews. *Oncology Nursing Forum, 42*(1), 89-93. doi: 10.1188/15.ONF.89-93

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA. Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. doi: 10.1371/journal.pmed.1000097

Monedero, J., Murphy, E. E., & O'Gorman, D. J. (2017). Energy expenditure and affect responses to different types of active video game and exercise. *PLoS ONE, 12*(5), e0176213. doi: 10.1371/journal.pone.0176213

Norris, L. A., Didymus, F. F., & Kaiseler, M. (2017). Stressors, coping, and well-being among sports coaches: A systematic review. *Psychology of Sport and Exercise, 33,* 93-112. doi: 10.1016/j.psychsport.2017.08.005

Park, J. M., & Noh, G. Y. (2017). Effects of the interactivity of mobile fitness application on exercise self-efficacy and attitude. *Advanced Science Letters, 23,* 1677-1681. doi: 10.1166/asl.2017.8622

Petosa, R. L., & Holtz, B. (2013). Flow for exercise adherence: Testing an intrinsic model of health behaviour. *American Journal of Health Education*, *44*(5), 273-277. doi: 10.1080/19325037.2013.811364

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences.* Oxford: Wiley-Blackwell.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M.,… Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews.* University of Lancaster, UK: ESRC Methods Programme.

Rheinberg, F. (2008). Intrinsic motivation and flow-experience. In H. Heckhausen, & J. Heckhausen (Eds.), *Motivation and action* (pp. 323-348). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511499821.014

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow Erlebens [Measuring flow experience]. In J. Stiensmeier-Pelster & F. Rheinberg (Eds.), *Diagnostik von Motivation und Selbstkonzept* (pp. 261–279). Göttingen: Hogrefe.

Robinson, J., Dixon, J., MacSween, A., van Schaik, P., & Martin, D. (2015). The effects of exergaming on balance, gait, technology acceptance and flow experience in people with multiple sclerosis: a randomised controlled trial. *BMC Sports Science, Medicine and Rehabilitation*, *7*(8), 1-12. doi: 10.1186/s13102-015-0001-1

Rohatgi, A. (2019). *WebPlotDigitizer (Version 4.2) [Computer software]*. Retrieved from: https://automeris.io/WebPlotDigitizer

Schüler, J., & Brunner, S. (2009). The rewarding effect of flow experience on performance in a marathon race. *Psychology of Sport and Exercise, 10*, 168-174.

Seligman, M., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, *55,* 5-14. doi: 10.1037/0003-066X.55.1.5

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*, 747-770. doi: 10.1146/annurev-psych-010418-102803

Sirriyeh, R., Lawton, R., Gardner, P., & Armitage, G. (2012). Reviewing studies with diverse designs: The development and evaluation of a new tool. *Journal of Evaluation in Clinical Practice, 18*(4), 746-752. doi: 10.1111/j.1365-2753.2011.01662.x.

Smith, B., & McGannon, K. R. (2018). Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. *International Review of Sport and Exercise Psychology*, *11*(1), 101-121. doi: 10.1080/1750984X.2017.1317357

Smith, M. F. (2018). *Research methods in sport*. London: SAGE.

Sone, T., Nakaya, N., Ohmori, K., Shimazu, T., Higashiguchi, M., Kakizaki, M.,... Tsuji, I. (2008). Sense of life worth living (ikigai) and mortality in Japan: Ohsaki study. *Psychosomatic Medicine, 70*(6), 709-715. doi: 10.1097/PSY.0b013e31817e7e64.

Swann, C., Crust, L., & Vella, S. A. (2017). New directions in the psychology of optimal performance in sport: flow and clutch states. *Current Opinions in Psychology*, *16*, 48-53. doi: 10.1016/j.copsyc.2017.03.032

Swann, C., Crust, L., Jackman, P., Vella, S. A., Allen, M. S., & Keegan, R. (2017b). Psychological states underlying excellent performance in sport: Toward an integrated model of flow and clutch states. *Journal of Applied Sport Psychology*, *29*(4), 375-401. doi: 10.1080/10413200.2016.1272650

Swann, C., Crust, L., Jackman, P., Vella, S., Allen, M., & Keegan, R. (2017a). Performing under pressure: *Exploring the psychological state of clutch performance in sport. Journal of Sport Sciences*, *35*(23), 2272-2280. doi: 10.1080/02640414.2016.1265661

Swann, C., Jackman, P. C., Schweickle, M., & Vella, S. (2019). Optimal experiences in exercise: A qualitative investigation of flow and clutch states. *Psychology of Sport and Exercise*, *40*, 87-98. doi: 10.1016/j.psychsport.2018.09.007

Swann, C., Keegan, R. J., Piggott, D., & Crust, L. (2012). A systematic review of the experience, occurrence and controllability of flow states in elite sport. *Psychology of Sport and Exercise, 13*(6), 807-819. doi: 10.1016/j.psychsport.2012.05.006

Swann, C., Keegan, R., Crust, L., & Piggott, D. (2016). Psychological states underlying excellent performance in professional golfers: “Letting it happen” vs. “making it happen”. *Psychology of Sport and Exercise, 23*, 101-113. doi: 10.1016/j.psychsport.2015.10.008

Swann, C., Piggott, D., Crust, L., Keegan, R., & Hemmings, B. (2015). Exploring the interactions underlying flow states: A connecting analysis of flow occurrence in

European Tour golfers. *Psychology of Sport and Exercise*, *16*, 60–69. doi:10.1016/j.psychsport.2014.09.007

Swann, C., Piggott, D., Schweickle, M., & Vella, S. (2018). A review of scientific progress in flow in sport and exercise: Normal science, a crisis point, and a progressive shift. *Journal of Applied Sport Psychology*, *30*(3), 249-271. doi: 10.1080/10413200.2018.1443525

Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *ACM Computers in Entertainment*, *3*(3), art. 3A.

Thin, A. G., Hansen, L., & McEachen, D. (2011). Flow experience and mood states while playing body movement-controlled video games. *Games and Culture*, *6*(5). 414-428. doi: 10.1177/1555412011402677

Vlachopoulos, S. P., Karageorghis, C. I., & Terry, P. C. (2000). Hierarchical confirmatory factor analysis of the Flow State Scale in exercise. *Journal of Sport Sciences, 18*(10), 815-823. doi: 10.1080/026404100419874

Wollseiffen, P., Schneider, S., Martin, L. A., Kerherve, H.A., Klein, T., & Solomon, C. (2016). The effect of 6 h of running on brain activity, mood, and cognitive performance. *Experimental Brain Research, 234*, 1829-1836. doi: 10.1007/s00221-016-4587-7

World Health Organisation. (2018). *Global action plan on physical activity 2018-2030: More active people for a healthier world.* Geneva, Switzerland: World Health Organisation.

Table 1: Study characteristics and quality of the included studies.

Authors (year of publication)	Participant information: Size (Male/ Female)	Participant information: Mean age (*SD*)	Activity	Design	Measurement of flow: Dispositional or state flow	Measurement of flow: Measure(s) and procedure	Key findings	Quality score[1]
Barry, van Schaik, MacSween, Dixon, and Martin (2016)	47 (27/20)[2]	33 (12)	Exergaming (balance training) and TGB exercise interventions	Quant/Between-subjects repeated measures	State	Completed the FSS after sessions at baseline (T1) and the end of a 4-week intervention (T2)	FSS subscale mean scores higher at T1 and T2 in exergamers. FSS subscale mean scores all higher at T2 compared to T1 in exergamers, and higher in seven subscales in the TGB exercise group at T2. Significant differences in the exergaming group compared to the TGB exercise group in six of the nine FSS subscales. No significant time-by-exercise type interaction effect found for any flow subscale.	52.38%
Bronner, Pinsker, and Noah (2013)	7 (4/3)	29 (9.34)	Lab-based exergaming (dance)	Quant/Within-subject	State	Completed the Video Game Training Effect Questionnaire immediately after each condition	Significant difference in game-flow intensity between games, but no significant post hoc pairwise comparisons. Moderate, positive association between flow intensity and both VO_2 and MET.	54.76%
Bronner, Pinsker, and Noah (2015)	7 (4/3)	29 (9)	Lab-based exergaming (dance)	Quant/Within-subject	State	Completed the Exer-game Questionnaire immediately after exposure to each condition	Significant difference in game-flow intensity between games, but no significant post hoc comparisons. Game-flow intensity significantly and positively correlated with MET and engagement. No significant correlation between flow and usability.	52.38%
Butzer, Ahmed, and Khalsa (2016; yoga)[3]	60 (25/35)	25.2 (3.2)	Yoga intervention	Quant/Between-subjects repeated measures	Dispositional	Completed the DFS-2 before (T1) and after (T2) an 8-week intervention	Mean scores for global dispositional flow and all DFS-2 subscales improved from T1 to T2.	71.43%
Elbe, Barene, Strahler, Krustrup, and Holtermann (2016; Zumba)[3]	33 (0/33)	46.3 (9.6)	Zumba intervention	Quant/ Longitudinal	State	Completed the FShS after baseline (T1), mid-intervention (T2), and late-intervention (T3) sessions in a 12-week intervention	Flow increased from T1 to T2 and from T2 to T3. 45% were participating in self-organized physical activity 18 weeks post-intervention.	61.90%
Elbe, Strahler, Krustrup, Wikman, and Stelter (2010; excluding football)[2]	20 (0/20)	36.9 (5.6)	Running intervention	Quant/Cross-sectional[4]	State	Completed the FShS after a single session midway through 12-week or 16-week interventions	No significant difference in flow between strength and interval training. No significant relationship between flow intensity in a mid-intervention session and post-intervention physiological improvements in male strength training or interval running groups.	66.67%
	8 (8/0)	36.2 (5.0)	Interval running intervention					
	10 (10/0)	31 (6.0)	Running intervention					
	8 (8/0)	37.3 (7.2)	Strength training intervention					

Ersöz and Eklund (2017)	251 (104/147)	Males: 23.57 (4.18) Females: 22.76 (3.96)	Exercise (running *n* = 76, weightlifting *n* = 66, cardiovascular exercise n = 77, unspecified exercise activity *n* = 32)	Quant/Cross-sectional	Dispositional	Completed the DFS-2 online	Significant multivariate effect for stages of change in exercise on the DFS-2 subscales. Significant, small-to-medium effect for CG, C, and AE, with higher scores in the action and maintenance stages than those in the preparation stage. DFS-2 subscales significantly and positively associated with intrinsic, identified, and introjected regulation and negatively related to external regulation and amotivation. No significant differences in the DFS-2 subscales between males and females.	59.92%
Ersöz, Altiparmak, and Hülya Aşçı (2016)	782 (369, 413)	21.87 (2.17)	Unspecified	Quant/Cross-sectional	Dispositional	Completed the Turkish DFS-2 after an exercise class	Significant differences between BMI groups in the DFS-2 subscales. Underweight participants reported higher dispositional flow than overweight and obese exercisers. No significant differences in the DFS-2 subscales between genders.	69.05%
Grove and Lewis (1996)	96 (30/66)	25.5 (9.33)	Circuit training	Quant/ Longitudinal	State	Completed an adapted PEQ during and immediately after two weekly sessions for six weeks	Significant two-way interaction for time of assessment x HS. Increases in flowlike state intensity from early to late assessments across the sample, but a significant increase was only apparent in high HS participants. Significant main effect of prior circuit training experience on flowlike state intensity. Greater flowlike state intensity in circuit trainers with > 6 months of experience. No significant relationship between changes in self-reported HR and flowlike state intensity between the early and late stages of sessions.	78.57%
Huang et al. (2018)	583 (362/221)	n/a[5]	Lab-based exergaming	Quant/Cross-sectional	State	Completed an adapted FS after a single session	Flow intensity significantly and positively related to perceived challenge and enjoyment. Flow intensity not significantly related to need for exercise or need for achievement.	66.67%
Iida and Oguma (2013)	469 (97/372)	67.3 (8.8)	Tai-chi classes	Quant/Cross-sectional	State	Completed an adapted JFSS-2 after a single class	Flow intensity significantly higher in those with 1-2, 4-5, 5-10, and ≥ 10 years of experience compared to those with < 1 year of experience. Statistically significant path from flow intensity to ikigai.	61.90%
Iida and Oguma (2014)	279 (not reported)	67.9 (7.9)	Tai chi classes	Quant/ Longitudinal	State	Completed the JFSS-2 after a baseline class (T1) and follow-up class one year later (T2)	Flow intensity significantly higher at T2 compared to T1 in the entire sample. Flow intensity at T2 significantly higher in non-experts but not in experts. Flow intensity at T1 positively predicted SoC at T2. Flow intensity at T1 significantly predicted SoC at T2 in experts, but not in non-experts. SoC at T1 did not significantly predict flow intensity at T2, regardless of experience.	73.81%

Karageorghis and Jones (2014 - quantitative)	22 (11/11)	Males: 19.6 (1.6) Females: 20.3 (1.6)	Lab-based treadmill walking/running	Quant/Within-subject and between-subjects	State	Completed the SFS after exposure to each condition	Significant two-way interaction of music tempo x gender, with a significant interaction for flow. Flow intensity significantly higher in the slow, medium, fast, and very fast music tempi conditions compared to the control condition.	83.33%
Karageorghis, Jones, and Stuart (2008)	29 (14/15)	Males: 20.7 (1.1) Females: 20.4 (1.4)	Lab-based treadmill walking	Quant/Within-subject	State	Completed the FSS-2 after exposure to each condition	Flow intensity significantly higher in the mixed, medium, and fast tempi groups compared to the control condition.	78.57%
Karageorghis, Vlachopoulos, and Terry (2000)	1231 (211/1014) (6 did not indicate)	31.43 years (9.13) (120 did not report age)	Aerobic dance exercise classes	Quant/Cross-sectional	State	Completed the FSS after a single class	Significant positive associations between flow intensity and post-exercise feelings of positive engagement, revitalisation, and tranquillity, and a significant, but weak, negative relationship between flow intensity and physical exhaustion. Flow intensity explained variance in positive engagement, revitalization, and tranquillity, but negligible variance in physical exhaustion.	80.95%
Kliem and Wiemeyer (2010)	22 (5/17)	47.36 (13.14)	Lab-based exergaming (balance) and TB training	Quant/Cross-sectional[4]	Dispositional	Responded to questions assessing game-flow after the final session of a 3-week intervention	No significant differences in game-flow intensity between TB training and exergaming groups.	38.10%
Lee et al. (2016)	25 (11/14)	36.4 (14.8)	Lab-based exergaming (yoga, strength, and balance)	Mixed method/Cross-sectional (Quant) and descriptive investigation (Qual)	State	Completed a Korean version of the FSS-2 and responded to closed-ended and open-ended (verbally) questions on the KUUEQ after a single session	Age, perceptions of pain, and physical dysfunction not significantly related to flow intensity. Compared to previously published exercise values (Jackson & Eklund, 2004), exergamers reported significantly higher C and AE. Level of difficulty was the most common reason for highest and lowest flow intensity. Excessive or insufficient difficulty was adversative, whereas optimal difficulty levels facilitated flow.	46.66%
Lee, Myers, Park, Hill, and Feltz (2018)	15 (10/5)	44.8 (9.72)	Home-based exergaming (cycling) and cycling intervention	Quant/Between-subjects and repeated measures	State	Completed an adapted FShS after sessions at days nine and 10 (T1), 41 and 46 (T2), and 65 and 70 (T3) in a 24-week intervention	Flow intensity mean scores increased over time in the entire sample Compared to the control condition, a NASP had a significant large effect on flow intensity at T2 and T3, while an ASP had a non-significant medium and negligible effect on flow intensity at T2 and T3, respectively.	88.10%
Loveday and Burgess (2017 – quantitative)	202 (74/124) (4 did not indicate)	29.3 (8.7)	Exergaming (mobile exergame)	Quant/Cross-sectional	State	Completed the SFS online	Flow intensity significantly predicted by game level, playing alone, nostalgia, and playing with family, but not by playing with friends or strangers.	45.24%
Marston, Kroll, Fink, and Gschwind (2016)	50 (23/27)[6]	n/a	Home-based exergaming (balance training) intervention	Quant/Cross-sectional[4]	State	Completed the AFSS after the final session of a 16-week intervention	No significant differences in global flow intensity or the nine AFSS subscales between age groups or study sites. Significantly higher mean scores for females on two of the 26 AFSS items. Significant differences in scores for 19 of the 26 AFSS items between study sites.	66.67%

Monedero, Murphy, and O'Gorman (2017)	23 (11/12)	24.8 (1)	Lab-based running and exergaming (ET-VG and FT-VG)	Quant/Within-subject	State	Completed the CFS after exposure to each condition	Significant effect of exercise type on core flow. Higher flow intensity for ET-VG, followed by FT-VG, self-selected intensity running, and running at 55%VO2Reserve.	50.00%
Park and Noh (2017)	100 (50/50)	-	Lab-based exergaming (unspecified exercise with mobile fitness applications)	Quant/Within-subject	State	Completed an unspecified measure after exposure to both conditions	Significant difference in flow intensity based on fitness application interactivity. Higher flow intensity reported while using an enhanced interactivity fitness application.	54.76%
Robinson, Dixon, MacSween, van Schaik, and Martin (2015)	56 (18/38)[7]	52 (5.8)	Lab-based exergaming (postural control) and TB training	Quant/Between-subjects and repeated measures	State	Completed the FSS after the first (T1) and final (T2) sessions of a 4-week intervention	Significantly higher scores in exergamers compared to the TB group in five FSS subscales at T2. Mean scores higher in T2 versus T1 for exergamers in all FSS subscales, but only higher in seven subscales for the TB training group at T2.	66.67%
Swann, Jackman, Schweickle, and Vella (2019)	18 (9/9)	32.94 (10.18)	Yoga, duathlon, running, treadmill running/walking, surfing, weight-training, hiking, half ironman, and indoor climbing	Qual/Descriptive investigation	State	Participated in event-focused, semi-structured interviews (*M* = 55 minutes) within two days of a self-reported rewarding exercise activity on average	Flow was characterised by 12 features: absence of negative thoughts; absorption; altered perceptions; automatic skill execution; confidence; ease/reduced effort; effortless attention; enjoyment; feeling in control; optimal arousal; motivation for more; and positive-in-the-moment feedback. Flow occurred in contexts that involved exploration, novelty/variation, and flexible outcomes through a gradual build-up of five stages: initial positive event; positive feedback; increase in confidence; challenge appraisal; and setting open goals. After flow, participants reported a sense of achievement, positive mood and emotions, intrinsic motivation, confidence, and energy.	79.48%
Thin, Hansen, and McEachen (2011)	14 (9/5)	19 (1.5)	Lab-based exergaming and cycling	Quant/Cross-sectional	State	Completed the FSS-2 after completing a single session	Compared to previously published values in exercise (Jackson & Eklund, 2004), exergamers reported significantly higher CSB and AA.	50.00%
Wollseiffen et al. (2016)	11 (6/5)	36.5 (7)	Outdoor running	Quant/ Longitudinal	State	Completed the SFS before and during a 6-hour run. Completed an adapted FSS-2 after the activity[8]	Significant changes in flow intensity over time in the run. Significant increase in flow intensity from pre-run to 1-hour, before a significant decrease from 1-hour to 3-hours, and declining thereafter. Flow intensity was higher in females before, during, and after the run, with significantly higher values at the pre-run, 3-hour, and 5-hour time points.	54.76%

Notes: (1) full details on the study quality assessment can be found in Supplementary File 4; (2) Barry et al. (2016) reported that data for 44 participants (exergame group *n* = 23 and TGB exercise group *n* = 21) were analysed due to data capture errors; (3) Some findings were not reported as the analyses included data for non-exercise participants; (4) Studies that involved exercise interventions but only collected data at a single point in time were categorised as cross-sectional designs; (5) no mean age was provided but the age ranges were as follows: 19-21 years old *n* = 285; 22-24 years old *n* = 283; 25-29 years old *n* = 14; missing *n* = 1; (6) Although a total of 153 participants were recruited, only 50 of the 95 exergame participants returned data on flow. The mean age for the sample (*n* = 153) was 76.15 years; (7) Data analysed for 46 participants due to the withdrawal of 10 participants; (8) Only data for items matching the SFS were analysed; (9) Abbreviations used as follows: AA = action-awareness merging; AE = autotelic experience; AFSS = Activity Flow State Scale; ASP = always superior partner; BMI = Body mass index; C = concentration on the task at hand; CFS = Core Flow Scale; CG = clear goals; CSB = challenge-skills balance; DFS-2 = Dispositional Flow Scale-2; ET-VG = entertainment themed-video game; FS = Flow Survey; FShS = Flow Short Scale; FSS = Flow State Scale; FSS-2 = Flow State Scale-2; FT-VG = Fitness themed-video game; HR = heart rate; HS = hypnotic susceptibility; JFSS-2 = Japanese Flow State Scale-2; KUUEQ = Korea University User Experience Questionnaire; MET = metabolic equivalent of the task; NASP = not always superior partner; PEQ = Privette Experience Questionnaire; Quant = quantitative; Qual = qualitative; SC = sense of control; SFS = Short Flow Scale; SoC = sense of coherence; T1 = time point 1; T2 = time point 2; T3 = time point 3; TB = traditional balance; TGB = traditional gym-based.

Table 2: Conceptual frameworks, item quantity, psychometric properties, result reporting levels, and issues with quantitative flow measures.

Conceptual framework	Measure	Study	Items	Validated measure of flow	Internal consistency reported (α)	Reporting Level		Issues[1]
						Subscale	Global	
Nine-dimensions framework	Activity Flow State Scale	Marston, Kroll, Fink, and Gschwind (2016)	26	Yes	No	Yes	Yes	E
	Dispositional Flow Scale-2	Butzer, Ahmed, and Khalsa (2016)	36	Yes	No	Yes	Yes	
		Ersöz and Eklund (2017)		Yes	Yes (.76 - .89)	Yes	No	A
	Flow State Scale	Barry, van Schaik, MacSween, Dixon, and Martin (2016)	36	Yes	No	Yes	No	A,C
		Karageorghis, Vlachopoulos, and Terry (2000)		Yes	Yes (.65-.84 for subscales)	No	Yes	B,C
		Robinson, Dixon, MacSween, van Schaik, and Martin (2015)		Yes	No	Yes	No	A,C
	Flow State Scale-2	Karageorghis, Jones, and Stuart (2008)	36	Yes	No	No	Yes	B
		Thin, Hansen, and McEachen (2011)		Yes	No	Yes	No	A
	Japanese Flow State Scale-2	Iida and Oguma (2014)	36	Yes	Yes (.91)	No	Yes	B
	Korean Flow State Scale-2	Lee et al. (2016)	36	Yes	No[2]	Yes	Yes	
	Short Flow State Scale-2	Karageorghis and Jones (2014)	9	Yes	No	No	Yes	B
		Wollseiffen et al. (2016)		Yes	No	No	Yes	B
		Loveday and Burgess (2017)		Yes	No	No	Yes	B
	Turkish Dispositional Flow Scale-2	Ersöz, Altiparmak, and Hülya Aşçı (2016)	36	Yes	Yes (.82)	No	Yes	B
	Adapted Flow State Scale-2	Wollseiffen et al. (2016)	9[3]	Yes	No	No	Yes	B,F
	Adapted Japanese Flow State Scale-2	Iida and Oguma (2013)	28[4]	No	Yes (.93)	No	Yes	B,F
	Adapted Flow Survey	Huang et al. (2018)	3	No	Yes (.86)	No	Yes	B,F
	Core Flow Scale	Monedero, Murphy, and O'Gorman (2017)	10	Yes	No	No	Yes	B
Game-flow	Video Game Training Effect Questionnaire	Bronner, Pinsker, and Noah (2013)	14	No	No	No	Yes	B,D
	Exer-game Questionnaire (game-flow subscale)	Bronner, Pinsker, and Noah (2015)	Unspecified	No	No	No	Yes	B,D
	Unspecified	Kliem and Wiemeyer (2010)	Unspecified	No	No	No	Yes	B,E
Eight-dimensions framework[5]	Flow Short Scale	Elbe, Barene, Strahler, Krustrup, and Holtermann (2016; Zumba)	10	Yes	Yes (.76)	No	Yes	B,G
		Elbe, Strahler, Krustrup, Wikman, and Stelter (2010; excluding football)	10	Yes	Yes (.75)	No	Yes	B,E,G
	Adapted Flow Short Scale	Lee, Myers, Park, Hill, and Feltz (2018)	8	No	Yes (.85)	No	Yes	B,F
Flowlike state	Adapted Privette Experience Questionnaire	Grove and Lewis (1996)	10	No	Yes (.81)	No	Yes	B,D
Unspecified	Unspecified	Park and Noh (2017)	5	No	Yes (.83)	No	Yes	B,D

Notes: (1) key issues that could not be detected by the quality assessment tool were as follows: (a) analysed flow at a subscale/dimensional level only; (b) analysed flow at a global level only; (c) measure used found to have inadequate model fit (Vlachopoulos, Karageorghis, & Terry, 2000); (d) measure used not previously validated; (e) longitudinal interventions that only assessed flow at a single time point; (f) removed items from a previously validated questionnaire; and (g) used cut-off points generated from data outside of exercise to classify the level of flow; (2) the internal consistency ($\alpha = .61$) of this translated questionnaire was below acceptable in previous research; (3) used two measures of flow and only analysed data in relation to the nine items on the Flow State Scale-2 that correspond with the Short Flow State Scale-2; (4) only seven of the nine dimensions were assessed; (5) the eight dimension framework (Rheinberg, 2008) includes all components of the nine-dimensions framework, with the exception of autotelic experience; (6) Author details and references for validated quantitative measures used by included studies are available in Supplementary File 5.

Figure 1: Flow diagram illustrating the screening process.