RESEARCH PAPER

Psychological Flow Scale (PFS): Development and Preliminary Validation of a New Flow Instrument that Measures the Core Experience of Flow to Reflect Recent Conceptual Advancements

Cameron Norsworthy[1] · James A. Dimmock[2] · Daniel J. Miller[2] · Amanda Krause[2] · Ben Jackson[1,3]

Abstract

In this study, we sought to develop—and provide preliminary validity evidence for scores derived from—a new Psychological Flow Scale (PFS). We propose a parsimonious model of three core dimensions of flow, reflecting the findings from a recent scoping review that synthesised flow research across scientific disciplines. The validation process for the PFS addressed recent conceptual criticisms of flow science regarding construct validity, theoretical compatibility, relational ambiguity, and definitional inconsistency. An initial review and analysis of the many flow measurements that exist found that these instruments either assess one, some, or none of the three core-dimensions of flow; often measuring similar dimensions that may bear resemblance to one of the three-dimensions but differ in dimensional meaning. PFS item development involved a phase of theoretical scrutiny, review of existing instruments, item generation, and expert review of items. Subsequently, 936 participants were recruited for scale development purposes, which included sample testing, exploratory factor analysis, and confirmatory factor analysis. This factor analytic process showed evidence for three distinguishable dimensions 'under' a single general or higher-order factor (i.e., global flow). With respect to external aspects of validity, flow scores correlated positively with perceptions of competence, self-rated performance, autotelic personality, and negatively with anxiety and stress scores. In conclusion, we present preliminary evidence for the theoretical and operational

Public Significance Statement This article presents evidence of validity for a new Psychological Flow Scale (PFS) that assesses the core aspects of the flow state, and has been designed to be utilized across domains and scientific disciplines. Confirmatory findings suggest that the response scale is a suitable fit to assess flow globally, whilst the dimensions of absorption, effort-less control, and intrinsic reward can be assessed by the three subscales. Flow scores correlated positively with perceptions of competence, self-rated performance, autotelic personality; and correlated negatively with anxiety and stress. The article offers suggestions regarding important aspects of flow measurement.

Extended author information available on the last page of the article

suitability of the PFS to assess the flow state across scientific disciplines and activity domains that be useful for experimental research and enable comparative flow research in the future.

1 Introduction

Flow, an optimal experience often colloquially referred to as being 'in the zone', has been studied extensively for more than 40 years because it is recognised as an important concept in the scholarly literature and popular culture. This psychological state of absorption and effortlessness to one's actions (Csikszentmihalyi, 1975) has historically been conceptualised through Csikszentmihalyi's model of flow, which contains nine key dimensions (see Jackson & Csikszentmihalyi, 1999). Continued debate regarding measurement and conceptual inconsistencies, however, has led to recent appraisals that "flow research is approaching a crisis point" (Swann et al., 2018; p. 249). Specifically, Swann et al. (2018) contended that Csikszentmihalyi's flow model is not sufficiently mechanistic to be considered a 'theory' as it lacks specific definitions and propositions underpinning testable causal relationships, and has issues of discriminant validity and conflation of flow with other states within the nine dimensions model. The nine-dimensional model has also been criticised regarding construct validity issues and relational ambiguity between dimensions (e.g., Engeser & Schiepe-Tiska, 2012; Heutte et al., 2021), theoretical incompatibility with other established psychological theories (e.g., self-efficacy; Jackson & Csikszentmihalyi, 1999; Swann et al., 2018), and inconsistent application and assessment across disciplines (e.g., Norsworthy et al., 2021; Peifer, 2012). Further, in experimental research the nine-dimensional model is often replaced by a unidimensional measure (challenge-skill balance) that (a) may not reflect the experience of flow, (b) is commonly accepted as an antecedent to flow rather than an experiential dimension, and (c) most commonly assess objectively through difficulty level with no subjective input (Norsworthy et al., 2021). In a recent collection of commentaries and articles on flow, Peifer and Engeser (2021a, b) concluded that "a gold standard for the modelling and measurement of flow is not at close reach" (p. 61). In reflection of recent research, Peifer et al. (2022) proposed a core experience of flow can be considered. Taken together, the critiques provided by these authors collectively highlight the need for researchers to re-consider issues of definition and conceptualization of flow, and then devise an (or highlight an existing) appropriate instrument that assesses any theoretical advancements regarding the flow construct.

In order to chart disparities and commonalities in contemporary flow research, Norsworthy et al. (2021) recently conducted a scoping review encompassing over 230 flow-related works spanning multiple scientific disciplines such as psychology, physiology, and neuroscience. Norsworthy and colleagues reported that flow was assessed using 141 different measures and described using 108 varying

constructs, terms, or dimensions—targeting all, some, or none of Csikszentmihalyi's nine dimensions (e.g., Fong et al., 2015; Zito et al., 2018). Within Norsworthy et al.'s synthesis, for instance, it was reported that Heutte et al. (2016) observed that barely half of the nine dimensions were perceived by learners in educational settings, and that specific dimensions—such as 'time transformation' or 'loss of self-consciousness'—are only considered relevant in specific contexts (e.g., Sinnamon et al., 2012; Swann et al., 2012). A common theme that emerged in Norsworthy and colleagues' review was the use of varied descriptive constructs that, despite similarities with one (or more) of Csikszentmihalyi's dimensions, were contributing to challenges when synthesising research findings (also see Auld, 2014). Examples of such instances included the use of 'effort-lessness' in psychophysiological settings (Bian et al., 2016; De Manzano et al., 2013) to characterise Csikszentmihalyi's 'sense of control'; 'telepresence' in human computer interface and gaming contexts (e.g., Klasen et al., 2012; Lazoc & Luiza, 2012) to characterise absorption or 'merging of action and awareness'; and the use of 'enjoyment' or 'intrinsic reward' (e.g., Llorens et al., 2013; Romero & Calvillo-Gamez, 2014) to represent 'autotelic experience'.

Based on their review findings, Norsworthy et al. (2021) concluded that despite substantial differences in terminology, flow researchers across scientific disciplines appeared to commonly conceptualise two key antecedents to flow, three core experiential dimensions, and three outcome themes. The antecedents were 'optimal challenge' which incorporated the concepts of clear task goals, immediate and unambiguous feedback; and 'high motivation' that subsumed the themes of interest, subjective value, intrinsic or extrinsic motives. The three core experiential dimensions that characterise the construct or experience itself were identified as 'absorption', 'effort-less control', and 'intrinsic reward'. Lastly, three common outcome themes of 'positive development', 'high functioning', and 'further engagement' were proposed. Similarly, in a recent literature review targeting theoretical integration in the field of flow research, Peifer and Engeser (2021b) expressed support for synthesising flow descriptors into three similar core experiential constructs (i.e., absorption, perceived demand-skill balance, enjoyment).

1.1 Three Core Experiential Dimensions to Flow

With respect to the absorption dimension of flow, Norsworthy et al. (2021) defined 'absorption' as "a state of absorption in the task characterised by focused, undistracted attention, and a merging of action and awareness". From a neuro-scientific perspective, Dietrich (2004) and Norsworthy et al. (2021) highlighted that flow is thought to occur through a depleting 'onion-peeling' effect (or down-regulation) of higher cognitive processes as attentional resources are reallocated to deal with the growing demands of the task (also see Sadlo, 2016; Ulrich et al., 2016). In line with Peifer and Engeser's (2021b) rationale, Norsworthy et al. outlined that as absorption occurs to meet the complexity of the task, a greater number of unnecessary higher cognitive functions (i.e., reflective self-consciousness

or time monitoring) are down-regulated to free up attentional capacity—facilitating frequent descriptions of flow nuances such 'loss of self-consciousness', 'time transformation', and a 'feeling of connection'. This delayering process may help to account for the seemingly continuous experience (i.e., mild to intense) of flow, after an initial threshold or discrete measure is passed (Norsworthy et al., 2021).

'Effort-less control' was defined by Norsworthy et al. (2021) as "a high sense of control in which the task feels less effortful than is typical for that person, characterised by fluidity of performance and an absence of concern over losing control". This dimension, which involves *less* effort (not void of all effort), or a sense of effortlessness, to one's high sense of control in the act, differentiates flow from other models of focused engagement that require high degrees of felt effort (i.e., cognitive/mental effort and high arousal) to override external distractors, and from other immersive states that may be enjoyable but not permit high functioning (Harris et al., 2017; Norsworthy et al., 2021; Peifer et al., 2014; Romero & Calvillo-Gamez, 2014; Tozman et al., 2015). The dimension 'effort-less control' differs from Peifer and Engeser's (2021b) utilisation of 'perceived demand-skill balance' on three key points. First, although Peifer and Engeser (2021b) explained that perceived demand-skill balance represents the high level of control felt in flow, sources within Norsworthy and colleagues review highlighted (in both psychological and neuroscientific research) that it is the sense of 'effortlessness' (i.e., a subjective sense of the act being less effortful or more fluid than usual) towards the sense of control that differentiates flow from other forms of high control (also see Peifer & Tan, 2021). Second, the optimal level of challenge (i.e., perceived demand-skill balance) is widely recognised (including by Peifer and Engeser) as an antecedent to, rather than a dimension of, flow (also see Barthelmäs & Keller, 2021). Lastly, Peifer and Engeser's (2021b) utilisation of 'perceived demand-skill balance' posits that flow must derive from a situation in which the individual's skill is being challenged. Norsworthy and colleague's utilisation of 'effort-less control', however, also accounts for the high degree of felt control in flow within non-demand-skill specific scenarios, such as non-achievement scenarios (e.g., an interesting conversation uninspired by achievement motives; see Engeser & Schiepe-Tiska, 2012).

The third dimension of flow, 'intrinsic reward', is characterised by positive valence and optimal levels of arousal (also see; Peifer et al., 2014; Ulrich et al., 2016). This dimension is evident in the activation of midbrain reward structures (Nah et al., 2017) and increased dopamine production that occurs during flow (Bian et al., 2016). The label 'intrinsic reward' captures the autotelic experience (as used by Csikszentmihalyi, 2014) and enjoyment (as used by Peifer & Engeser, 2021b) of flow and was chosen by Norsworthy et al. (2021) because the term better represents what was observed in the literature—in fact, Csikszentmihalyi himself, often utilises the term 'intrinsic reward'. 'Intrinsic reward' is more widely applicable across scientific disciplines and can be assessed physiologically (e.g., dopamine levels) without involving reflective cognitive processes that occur following the flow experience, as would be necessary for determining one's level of enjoyment—opening up the possibilities for potential bias from outcomes and contextual or social factors (Abuhamdeh, 2021). For further details, see Norsworthy et al.'s (2021) and Abuhamdeh's (2021) reviews on the relationship between flow and enjoyment.

Given the scope and findings of their review (which included an examination of existing flow instruments), Norsworthy et al. (2021) suggested that no existing flow instrument adequately assesses this three-dimensional conceptualisation of flow, and that a new flow instrument was required to assess these dimensions and exact conceptualisation of flow. Although many flow measurements exist, they either assess one, some, or none of the three-dimensions; often assessing similar dimensions (e.g., enjoyment) that may bear resemblance to one of the three-dimensions (e.g., intrinsic reward) but differ in dimensional meaning. The main purpose of this study, therefore, was to examine existing instruments, develop an instrument that could assess flow experiences using this three-dimensional conceptualisation of flow, and to examine aspects of the reliability and validity of scores derived from that instrument. Instrument development followed Boateng et al.'s (2018) three-phased approach (i.e., item development, scale development, scale evaluation) and was grounded in Messick's (1995) recommendations for the assessment of different aspects of construct validity. Specifically, we sought to provide preliminary evidence for content, substantive, structural, external, and generalizability aspects of validity; content validity: relevance, representativeness, technical quality; substantive validity: a theoretical rationale; structural validity: factor structure; external validity: convergent and discriminant evidence; and generalizability validity: examining scores across populations, settings, and tasks aspects of validity.

1.2 Transparency and Openness

We describe our sampling plan, all data exclusions (if any), all manipulations, and all measures in the study, and the Transparency and Openness Promotion (TOP) Guidelines (Nosek et al., 2015) were adhered to. All data, analysis syntax, and research materials can be found in the article, S1, and the Appendix 1. Data analysis is detailed below. The study protocol was approved (RA/4/20/609) by the lead author's institutional ethics committee prior to data collection.

2 Phase 1: Item and Scale Development

We used a multi-stage approach for item development including theoretical consultation, review of existing instruments, item generation, item review, expert review of items, and target population review of items. Norsworthy et al.'s (2021) scoping review findings (see Table 1) provided the conceptual justification for item generation, and supported the substantive aspect of validity for instrument development. Although antecedent constructs have been assessed within existing flow measures (e.g., Flow Short Scale, Rheinberg et al., 2002), Norsworthy et al.'s proposed antecedents to flow (i.e., 'optimal challenge', 'high motivation') were excluded from item development to ensure the new instrument assessed the flow experience itself, and did not conflate experience with putative antecedents or pre-conditions.

2.1 Examining Existing Instruments[1]

An initial search within the databases Web of Science (1900), PubMed (1997), Medline (1966), Scopus (1966), Embase (1947), PsycINFO (1806), SPORTDiscus (1930), and Google Scholar was conducted to ensure no existing instrument adequately assessed these three dimensions. No time parameters were attributed. Many flow assessments exist (including over 20 Likert scale instruments)*, but they either assess one, some, or none of the three-dimensions (absorption, effort-less control, intrinsic reward); often assessing similar dimensions (e.g., high control or enjoyment) that may bear resemblance to one of the three-dimensions (e.g., effort-less control or intrinsic reward) but differ in dimensional meaning. The widely used FSS-2 (Jackson & Eklund, 2002) assessment, for example, assesses all nine-dimensions to flow; not only does this assessment include the antecedents of optimal challenge (and sub-dimensions: clear goals and unambiguous feedback), it also does not directly assess effort-less control and assesses the nuanced dimension of 'time transformation' that is often criticised for its validity. The Autotelic Personality Questionnaire (APQ; Tse et al., 2020) assesses personality traits and not the flow experience, and many of the

---

[1] Measures found included (not the exhaustive list): Activity Flow Scale; adrenocorticotropic hormone [ACTH]; Autotelic Personality Questionnaire (APQ); blood pressure (BP); Blood Volume Pulse (BVP); cardiac output; cortisol levels; challenge-skill - 'flow-simplex'; CSBI - Challenge-Skill Balance Index; Flow Questionnaire (Flow Q); Deep Structured Experiences Questionnaire; Dispositional Flow Scale (DFS, DFS short, DFS-2, & DFS-2 short); Day Reconstruction Method (DRM); ECG-cardiac output; ECG-Total Peripheral Resistance (TPR); ECG-cortical activity; ECG-HRV; Electrodermal Activity (EDA); EduFlow Scale; EEG-affective states; EEG-emotional valence; EGameFlow; EMG-corrugator supercilia; EMG-orbicularis oculi; EMG-zygomaticus major; EMG- IM and sensorimotor behaviour; EMG- motor commands; EMG-EEG for CMC cortico-muscular coherence; endocrinological parameters; epinephrine levels; Experience sampling method-FQ; multiple ESM-multicomponential/dimensional measures; eye blinking; eye pupil activity, eye tracking; facial expressions; Flow in Environment Scale (FLIES); Flow experience scale by Shin (2006); Flow Contextual Questionnaire (CFC); Flow Metacognitions Questionnaire (FMQ); Flow Questionnaire; Flow Scale; Flow Short-Scale (FSS); Flow Synchronization Questionnaire (FSyQ); FLOW-W Questionnaire; Functional Magnetic Resonance Imaging (fMRI)-decreased left amygdala activity; fMRI-continuous arterial spin labelling; Functional Magnetic Resonance Perfusion Imaging (fMRPI); Functional Near-Infrared Spectroscopy (fNIRS); Flow Quadrant; Flow State Scale (FSS, FSS short, FSS-2, & FSS-2 short); and multiple adaptations to the FSS; Galvanic skin response (GSR); Heart Period (HP); High Frequency (HF) band; Hypothalamic-Pituitary-Adrenal (HPA) levels; Heart Rate (HR); HR coherence; Heart Rate Variability (HRV); Impedance-Cardiographic Signals (ICG) levels; Interbeat Interval (IBI); Internet Flow Scale (IFS); multiple interview methods; kinematic motion segments; Low Frequency (LF) band; LF/HF ratio; mental effort measures; Magnetic Resonance (MR) activity; NIRS levels-oxygenated hemoglobin (O2Hb) concentrations; Observation methods; Operant Motive Test (OMT); Oxygenation levels (using near infrared spectroscopy); Oxytocin and ACTH concentrations; Positron Emission Tomography (PET)-default mode network; PET-dopamine-D2R binding potential (BPND); PET scan with C-labeled raclopride RAC; Phasic EDA; Skin Conductance Response (SCR); PsychoLog (phone app); Relational Flow Scale (RFS); respitory rate; Respiration Signal (RSP)- salivary cortisol; self-report sliders; State Flow Scale (SFS); Short Flow in Work Scale (SFWS); Skin Conductance (SC); Skin Resistance (SR); Social Network Analysis (SNA); Stroke volume (SV); sustained attention response; Stroke Volume (SV) blood measure; Swedish Flow Proneness Questionnaire (SFPQ); Team Flow Monitor TFM; the experience fluctuation model; the zone test; Tonic EDA-long-term skin conductance level (SCL); top down attention: Gaze variability; total peripheral resistance; Transcutaneous (through the skin) Vagus Nerve Stimulation (tVNS); ventricular contractility; Work-related Flow Scale (WOLF, I-WOLF, & WOLF-S); and over 30 custom scales. For a full list of these measures, what they mean, the dimensions that they measure, and references of these measures, see Norsworthy et al. (2021).

physiological assessments mentioned above are still elementary and do not distinguish as to whether they assess a specific dimension or more accurately assess the experience as a whole. The Flow Short Scale (Rheinberg et al., 2002) was considered the closest existing representative instrument (to the three-dimensions posited in Norsworthy et al.'s scoping review; see Norsworthy et al., 2021, Chap. 2), due to the measures of fluency of performance and absorption. The FSS consists of 13 items that are suggested to predominantly assess the two factors of 'fluency of performance' and 'absorption' (whilst also offering questions to assess 'perceived importance', 'perceived outcome importance', 'demand', 'skills', and 'perceived fit of demands and skills') resembling two of the three dimensions targeted in this study (i.e., absorption, effort-less control). The FSS however, (a) did not fully assess the dimensions (and sub-dimensions) of absorption and effort-less control as laid out in the scoping review, (b) includes antecedent constructs, and (c) does not assess the intrinsic reward dimension. Although the existing options for assessing flow are plentiful, and utilising (or developing) an existing measure would certainly be advantageous, when examining existing instruments to assessing flow it was clear that no existing measurement instruments adequately assesses the core conceptualisation of flow that Norsworthy et al. (2021) highlighted in the recent scoping review. The existing measures have substantiative issues such as confounding antecedents into the experience assessment and not directly assessing the *effortlessnes*s associated with the high degrees of control that specifically delineates, and not conflates, flow from other similar states such as clutch states, for example.

2.2 Item Generation

Given that no adequate existing instrument was available, the lead author initially generated a pool of 60 potential items (20 for each dimension) to assess the three

Table 1 Norsworthy et al.'s (2021) Flow definition, dimensions, and descriptions

Proposed definition of flow:	Flow is an intrinsically rewarding state of absorption in a task in which a high degree of control feels effort-less
Antecedents to flow	Antecedents to Flow
Optimal challenge	A perceived capability to meet the challenging demands of the situation
High motivation	A high motivational force
Experiential dimensions to flow	
Absorption	A state of absorption in the task characterised by focused, undistracted attention, and a merging of action and awareness
Effort-less control	A high sense of control in which the task feels less effortful than is typical for that person, characterised by fluidity of performance and an absence of concern over losing control
Intrinsic reward	An intrinsically rewarding experience characterised by positive valence and optimal levels of arousal

Note: The above table and flowdimensions and descriptions has been taken from Norsworthy et al.'s (2021) scoping review

flow experience dimensions. These items were generated from dimensional and sub-dimensional descriptions and adapting questions from existing scales that targeted the same (or similar) dimensions. Two co-authors—both of whom had extensive experience with instrument development procedures—then provided detailed review and comments (e.g., identifying redundancy, ambiguity, and double-barrelled questioning) that resulted in a revised parsimonious pool of 36 potential items (12 for each dimension).

2.3 Expert and Target Population Testing

The initial pool of 36 items was first assessed by an expert panel comprising five scholars with research expertise in the topic (Boateng et al., 2018). These experts were chosen for their relevant academic experience evident by continuous publications in the motivational and psychological literature, and recruited through direct contact and the authors' academic network. They were asked to provide feedback on item content with an emphasis on relevance and technical quality (e.g., representativeness, understanding, jargon, overlap, ambiguity; Haynes et al., 1995). The experts were presented with all items alongside definitions of the three dimensions (which were used as criteria for the rating). They were then asked to provide a rating from 1 (*not representative*) to 5 (*clearly representative*) in terms of how well each item represented the targeted dimension. Additionally, we requested qualitative feedback on item clarity, wording, and possible dimensional overlap. Surveys were administered through the Qualtrics online platform and included written instructions and definitions of key terms prior to commencement.

In order to gather target participant (and also to serve as a 'non-academic' form of) feedback, a group of 15 adults from the general population (6 males, 9 females) were also recruited through the lead author's network. The participants came from a range of professions (e.g., student, teacher, executive, designer, sports coach, and retail). These participants were asked to undertake an identical task. Together, these consultations resulted in 8 items being dropped and minor wording changes to ensure item clarity and relevance. For example, items with a mean score below 4.4 were cut, and items containing the word 'performance' or terms deemed 'too academic' were removed or changed. As a result of these review stages, the initial pool of 60 items had been reduced to 28 items. These items are presented in the Supplementary Material (Table S1; S = supplementary material).

3 Phase 2: Evaluation Through Factor Analysis and Validity Testing

3.1 Overview and Procedure

We began Phase 2 with these 28 items and sought to test and further refine the item pool through the recruitment of a large sample of participants and iterative factor analytic methods. Our primary aim in Phase 2 was to perform exploratory and confirmatory factor analyses with the goal of testing and further refining the item pool

and instrument. We also sought to (a) examine aggregate- and item-level descriptive statistics, and (b) assess external aspects of validity by examining item responses against related constructs and general appraisals of flow.

We used a 7-point Likert-type response scale for the flow items anchored at 1 (*Strongly Disagree*) and 7 (*Strongly Agree*). This response scale was selected because there is evidence that scale reliability and validity are generally improved using a 7-point scale (Dawes, 2008). In addition to the flow items, we assessed related constructs and included questions assessing general appraisals of flow (see Stage 3 for full details). We collected data through the use of a cross-sectional survey with 913 participants (from an initial pool of 956 of which 43, 4.5%, had large sections of incomplete data). This sample consisted of adults (635 females, 285 males, 3 prefer not to say) from English-speaking countries (United Kingdom, United States, Australia, New Zealand, Ireland, South Africa), and all participants were recruited through the Prolific data collection platform. Prolific is a research recruitment and data collection company designed for use by certified academic institutions and only for research approved by an institutional review board. Prolific samples are increasingly prevalent in research reports, especially in cases when online methods are used for data collection within behavioural sciences (see Buhrmester et al., 2018; Palan & Schitter, 2018); there is evidence that online methods may reduce sample biases in comparison to traditional data collection approaches (e.g., Gosling et al., 2004). To ensure adherence and good quality data, time limits were set, participants were paid above average (see prolific) rates, and two attention tests were included in the questions (Aguinis et al., 2021).

All participants were provided electronically with an information letter and provided informed consent prior to completing the questionnaire. Data were downloaded and stored in a de-identified spreadsheet on a secure server by the first author to ensure participant confidentiality. Participants were instructed to engage in an activity of their choice before commencing survey participation—the questionnaire was (requested to be) completed immediately after activity participation to ensure experience recall was as close to the event as possible (as recommended for flow measures; Norsworthy et al., 2017). Participants were instructed that they could use an activity that they had already completed prior to filling in the questionnaire (within the last hour), and if no activity had been carried out then to stop and engage in an activity prior to continuing form completion. Participants consented to having carried out the activity. Items in the survey were focused on participants' thoughts and feelings experienced during the focal activity. Data from approximately half of the total sample ($n=452$) were randomly apportioned for exploratory factor analytic purposes (see Stage 1 below) to refine the item pool and instrument, and the remaining ($n=461$) data were assigned for confirmatory factor analytic purposes (see Stage 2 below). During analysis, data were initially screened for missing values. Pairwise deletion of cases was carried out for the few ($<0.01\%$) missing data items during the factor analysis. Finally, to assess discriminant and convergent validity, we used all participant data ($n=913$) to examine correlations between flow scores—using *only* the subscale and global scores that were ultimately retained for the 'final' instrument following stages 1 and 2—and related variables (see Stage 3 below). We felt that (re-)using the entire sample for this purpose was likely to generate more

representative correlation values (rather than, for example, computing these correlations separately with the sub-samples used for stage 1 and stage 2 analyses).

3.2 Stage 1: Exploratory Factor Analysis

3.2.1 Method

Exploratory factor analysis (EFA) was performed in Stage 1 to (a) explore the factor structure of items, (b) examine the structural aspect of validity (i.e., examine the higher-order structure of the item pool), (c) eliminate problematic (i.e., high cross-loading) items, and (d) identify the most parsimonious instrument (Boateng et al., 2018). Specifically, a principal axis factor analysis with a promax rotation was carried out. Promax rotation was chosen as it is an oblique rotation method. The analysis was performed using the EFA.dimensions package for R (version 0.1.7.6; O'Connor, 2023). This package was also used to determine the number of factors to extract. Eight such tests were applied: the empirical Kaiser criterion, comparison data, the Hull method, Velicer's minimum average partial test, parallel analysis, the salient loadings criterion, the standard error scree test, and the sequential chi-square model test. Parallel analysis was run using a principal components extraction and 95% percentile parallel eigenvalues, as these specifications show the highest accuracy across varied data conditions (Auerswald & Moshagen, 2019).

3.2.2 Results

A KMO value greater than 0.6 indicates factor analysis is appropriate, and greater than 0.9 is 'marvelous' (Kaiser, 1974; Kaiser & Rice, 1974). A significant Bartlett test ($p < .05$) is indicative of adequate conditions to fit a factor analysis. A strong KMO value (.92) and significant Bartlett's test of sphericity ($\chi^2 (378) = 8424.89$, $p < .001$) indicated adequacy of fit for factor analysis. Six of the eight factor extraction tests indicated that four was the optimal number of factors to extract. The standard error scree test and the sequential chi-square model test suggested extracting 7 and 12 factors, respectively. When there is an incongruence between the results of the sequential chi-square model test and parallel analysis, Auerswald and Moshagen (2019, p. 487) recommend referring to comparison data or the empirical Kaiser criterion (both of which suggested a four-factor solution). Accordingly, four factors were extracted. Figure S1 presents a plot with the observed and 95th percentile parallel eigenvalues produced as part of the parallel analysis.

Together, the four factors explained 63% of the variance in the data. Eigenvalues and the variance explained per factor are presented in Table 2. Factor loadings and communalities are also presented in this table. Based on consideration of item content, the factors were labelled (1) absorption, (2) effort-less control, (3) intrinsic reward, and (4) intuiting. This fourth factor was unexpected based on the theoretical model on which scale items were constructed.

With the goal of developing a parsimonious (final) instrument, the strongest 3 items across each of the four factors (12 items in total) were retained for confirmatory analysis (Stage 2). Items were retained based on high factor loadings (with loadings ≥ 0.45, 0.55, 0.63, and 0.71 being considered fair, good, very good, and excellent, respectively; Tabachnick & Fidell, 2013), low cross-loadings (≤ 0.32; Tabachnick & Fidell, 2013), and expert judgement as to whether items adequately covered all aspect of each construct. For example, Item 14 showed a lower factor loading but was retained for the CFA as the item description (i.e., letting it happen rather than making it happen) has been used as a key differentiating description between flow and other similar states (Swann et al., 2016).

3.3 Stage 2: Confirmatory Factor Analysis

3.3.1 Method

In Stage 2, confirmatory factor analysis (CFA) was conducted on the proposed 12 Psychological Flow Scale (PFS) items. Norsworthy et al.'s (2021) scoping review highlighted that, conceptually speaking, the flow experience should not be reduced solely to its dimensions, or be represented by only one or some of its dimensions (as researchers have done in the past); rather, flow is posited to occur when all three dimensions are present and interactive. Empirical testing to ascertain whether the PFS could include a single global latent factor (i.e., global flow score, as opposed to the instrument being limited to the assessment of the factors/dimensions), therefore, was deemed important. Accordingly, a series of models were tested: (1) a global latent factor model in which items were treated as reflective indicators of a single global flow factor, (2) a four factor, orthogonal model in which items were treated as reflective indicators of their purported factors as indicated by the EFA (absorption, effort-less control, intrinsic reward, and intuiting), but with no global flow factor, and (3) a higher-order model in which items were treated as reflective indicators of one of four first-order factors (absorption, effort-less control, intrinsic reward, and intuiting), which were themselves treated as reflective indicators of a second-order factor representing global flow. Figures depicting these models are presented in the supplementary material (Figure S2-S5). Model identification was achieved by setting the variances of latent factors to one. The analysis was carried out in Amos (version 27) using maximum likelihood estimation. When using maximum likelihood estimation, it is recommended that |skew| < 2 and |kurtosis| < 7 for all items (Fabrigar et al., 1999). All items showed skew and kurtosis below this threshold (see Table S1).

3.3.2 Results

As can be seen in Table 3, the single global factor model fit the data poorly, suggesting that the scale does not measure one domain only (Dunn & McCray, 2020). The four orthogonal factors model also showed poor fit to the data. However, loadings of factors on items were generally high (with the exception of item 14, all standardized

factors loadings were above 0.71; see Table S2 for factor loadings for Models 1–3). This would suggest that, although items are good indicators of their factors, the scale does not simply measure a series of lower-order dimensions. The higher-order model with one second-order global flow factor and four first-order factors showed marginal fit to the data. Large standardized factor loadings were observed between first-order factors and their purported items (again with the exception of item 14) and between flow and three of the four first-order factors. Flow showed a weak standardized loading on intuiting (0.32). Consulting the modification indices indicated that freeing the error terms for the intuiting and effort-less control factors to covary would reduce model discrepancy by 56.92. Freeing these error terms to covary resulted in the standardized factor loading from flow to intuiting to decrease substantially (from 0.32 to 0.14). We interpreted this to indicate that the intuiting factor was not caused by the global flow factor, but rather only appeared to be caused by flow due to shared variance with effort-less control. The poor factor loading from flow to intuiting, and the fact that an intuiting factor was unexpected based on theoretical understandings of flow (Norsworthy et al., 2021), prompted us to specify a fourth model excluding the intuiting factor and its items. This modified model was found to fit the data well, with the exception of the model χ^2 test (which is to be expected given the large sample size; Hoyle, 2011). Large standardized factor loadings (all >0.71; see Table 4) were observed between each first-order factor and its item indicators. Standardized factor loadings between the general flow factor and each second-order factor were all in excess of 0.52. This final model best reflects the theoretical model on which the scale items were constructed. The final model was also rerun with the full dataset ($N=913$), producing similar results (reported in Tables S3 and S4).

Internal consistency estimates (Cronbach's α and McDonald's ω) where then calculated for global flow, absorption, effort-less control and intrinsic reward on the full dataset. In all cases, these values (reported in Table 5) exceeded 0.82, indicating good internal consistency. Table 5 also reports descriptive statistics and intercorrelations between subscales.

3.4 Stage 3: Discriminant and Convergent Validity

In the final stage, we examined external validity (i.e., discriminant and criterion relevance) for scores derived using the final, 9-item version of the PFS through correlational analysis.

3.4.1 Method

To examine discriminant and convergent validity, PFS scores were correlated with constructs that have been previously linked (both positively and negatively) to flow. These constructs included cognitive state anxiety and pressure / tension (e.g., Llorens & Salanova, 2017; Peifer, 2012; Sadlo, 2016), perceived competence (e.g., Schuler & Brandstatter, 2013; Valenzuela et al., 2018), performance (e.g.,

Table 2 Promax rotated factor loadings, extraction communalities (h^2), proportion of variance explained by retained factors, and rotated factor correlations

Original item number	PFS item number	Factor loading				h^2
		Factor 1	Factor 2	Factor 3	Factor 4	
1		**-0.83**	0.00	0.03	-0.02	0.71
2	PFS1	**-0.83**	0.01	0.02	-0.03	0.72
3		**-0.70**	0.03	0.09	0.04	0.56
4	PFS2	**-0.89**	-0.04	-0.02	-0.04	0.78
5		**-0.87**	-0.06	0.02	0.05	0.73
6	PFS3	**-0.90**	-0.03	-0.04	0.03	0.76
7	†	-0.04	0.19	-0.01	**0.60**	0.45
8		**-0.48**	0.10	-0.15	-0.33	0.35
9	†	0.03	0.12	0.09	**0.56**	0.37
10		**-0.65**	0.06	-0.01	0.18	0.44
11		**-0.59**	0.12	-0.04	-0.14	0.41
12		0.20	**0.65**	-0.08	-0.14	0.36
13		0.10	**0.59**	0.09	0.22	0.49
14	†	0.00	0.23	-0.07	**0.42**	0.26
15	PFS4	0.06	**0.62**	0.01	-0.07	0.36
16	PFS5	-0.02	**0.87**	-0.15	0.18	0.76
17	PFS6	-0.11	**0.69**	0.01	0.23	0.64
18		-0.10	**0.54**	0.03	0.27	0.47
19		0.08	**0.63**	-0.02	**0.33**	0.59
20	PFS7	-0.02	-0.04	**0.89**	-0.01	0.78
21	PFS8	-0.03	-0.05	**0.92**	-0.04	0.83
22	PFS9	0.08	0.00	**0.90**	0.01	0.75
23		0.02	0.24	**0.54**	-0.27	0.51
24		-0.02	-0.07	**0.91**	0.10	0.79
25		-0.02	-0.14	**0.89**	0.06	0.70

Table 2 (continued)

26	0.01	0.20	-0.20	0.56
27	-0.07	-0.06	0.20	0.61
28	-0.47	0.03	0.05	0.28

	Initial values			After extraction		
Factor	Eigenvalue	% of Variance	Cumulative %	Sums of squared loading	% of variance	Cumulative %
1	8.91	0.32	0.32	8.57	0.31	0.31
2	4.54	0.16	0.48	4.10	0.15	0.45
3	2.78	0.10	0.58	2.44	0.09	0.54
4	1.48	0.05	0.63	0.91	0.03	0.57

Rotated factor correlations

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	-			
Factor 2	-0.22	-		
Factor 3	-0.46	0.45	-	
Factor 4	-0.16	0.21	-0.03	-

Note. Factor loadings taken from pattern matrix. Factor loadings greater than |0.32| are in bold. Extraction method = principal axis factoring

† indicates that item was included for CFA testing, but not retained in the final scale

Csikszentmihalyi et al., 2018; Flett, 2015; Moran, 2012), and autotelic personality (Baumann, 2012; Jackson & Csikszentmihalyi, 1999; Tse et al., 2020). Additionally, the Flow Short Scale (Rheinberg et al., 2002) was also included to evaluate scale redundancy as it was considered the closest existing flow scale (in terms of dimensional similarity) to the PFS. Evidence for generalisability (i.e., across activities) was considered inherent within the dataset because participants carried out a wide variety of activities (see Results). Common criticisms of existing psychological measurements of flow (see Ellis et al., 2019; Engeser, 2012; Jackman et al., 2017; Moneta, 2012; Swann et al., 2018) were also addressed through assessing a number of specific indicators of the flow experience. Specifically, participants were given direct questions (see 'Measures' below) targeting self-reports of flow entry, frequency of perceived flow entry and exit within the same event (to account for multiple occurrences of flow, or not), flow intensity (to examine intra-differences), and percentage of time in flow within the given event.

Assessments

Psychological Flow Scale (PFS) The 'final' PFS (for the purpose of this study, as confirmed during stage 1 and 2) consisted of nine items, with three items for each of the three flow dimensions (the items and the dimensions they pertain to are detailed in the Appendix 1). A global flow score was determined by averaging responses to the nine items. Subscale scores (i.e., absorption, effort-less control, intrinsic reward) were determined by averaging responses to the three items in each subscale.

Perceived Anxiety The Mental Readiness Form-3 (MRF-3; Krane, 1994) was utilised to understand participants' perceived level of anxiety during the task in question. The MRF-3 assesses cognitive state anxiety and was developed as a shorter alternative to the Competitive State Anxiety Inventory-2 (CSAI-2; Martens et al., 1990). The assessment involves one question, "Please rate how you felt during your chosen task/activity", in which participants respond on a response scale from 1 (*not worried*) to 11 (*very worried*).

Perceived Competence The Intrinsic Motivation Inventory (IMI) is a multidimensional instrument intended to assess participants' subjective experience related to a target activity. The 6-item perceived competence subscale of the IMI (Ryan, 1982) was used in this study to understand participants' perceived competence (e.g., "I think I am pretty good at this activity"). Responses were made on a response scale anchored at 1 (*not at all true*) and 7 (*very true*). Responses to all six items were summed to determine a score for further analyses. The internal consistency estimate for scores derived from this assessment in this study was $\alpha = 0.88$.

Perceived Stress The 5-item 'pressure/tension' subscale of the IMI (Ryan, 1982) was used to assess participants' stress perceptions (e.g., "I felt very tense while doing this activity"). Responses were scored on the same 7-point response scale

Table 3 Fit indices for tested CFA models

Fit measure	Model 1: Single global factor	Model 2: Four orthogonal factors	Model 3: Higher-order model (four first-order factors)	Model 4: Modified higher-order model (three first-order factors)
χ^2 test	$\chi^2 (54) = 1499.88, p < .001$	$\chi^2 (54) = 408.68, p < .001$	$\chi^2 (50) = 190.03, p < .001$	$\chi^2 (24) = 70.09, p < .001$
CFI	0.532	0.885	0.955	0.983
TLI	0.428	0.860	0.940	0.974
RMSEA	0.241	0.119	0.078	0.065
RMSEA 90% CI	0.231, 0.252	0.109, 0.130	0.066, 0.090	0.047, 0.082
SRMR	0.181	0.219	0.080	0.039
AIC	1547.88	456.68	246.03	112.09
BIC	1647.08	555.88	247.65	198.98

Table 4 Factor loadings for final model (Modified higher-order model with three first-order factors)

Factor loading	Unstandardized	SE	Standardized
Flow → Absorption	1.08***	0.17	0.73
Flow → Effort-less Control	0.63***	0.08	0.53
Flow → Intrinsic Reward	1.04***	0.16	0.72
Absorption→ Item 2	0.59***	0.06	0.78
Absorption → Item 4	0.72***	0.06	0.90
Absorption → Item 6	0.78***	0.07	0.79
Effort-less Control → Item 15	0.68***	0.05	0.72
Effort-less Control → Item 16	0.99***	0.05	0.88
Effort-less Control → Item 17	0.88***	0.05	0.86
Intrinsic Reward → Item 20	0.83***	0.07	0.87
Intrinsic Reward → Item 21	0.88***	0.07	0.95
Intrinsic Reward → Item 22	0.88***	0.07	0.89

Note. Item numbering in table refers to the initial 28 items

** $p < .01$; *** $p < .001$

described in the previous section. Responses to all five items were summed to determine a score for further analyses. The internal consistency estimate for scores derived from this assessment in this study was $\alpha = 0.82$.

Self-Reported Performance A single self-report item was used to examine perceived performance. Participants responded on a response scale ranging from 1 (*very low performance*) to 11 (*very high performance*) to the item, "Please rate how you felt you performed in your chosen activity".

Autotelic Personality The Autotelic Personality Scale (APQ; Tse et al., 2020) was used to assess autotelic disposition. The scale consists of 26 items including 7 sub-scales (curiosity, persistence, low-self-centeredness, intrinsic motivation, enjoyment and transformation of challenge, enjoyment and transformation of boredom, attentional control). Example items include "I am curious about the world" (curiosity), "I find it hard to choose where my attention goes" (attentional control), and "I am good at finishing projects" (persistence). Participants used on a response scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Responses to all 26 items were summed to determine a composite score for further analyses (Tse et al., 2021). The internal consistency estimate for scores derived from this assessment in this study was $\alpha = 0.84$.

Flow Short Scale (FSS) The FSS (Rheinberg et al., 2002) was used because it was deemed a common existing assessment of flow that most closely resembled the PFS. Consistent with Kyriazos (2018), only the first 10 items (out of 13)—which target the flow experience of 'fluency of performance' (e.g., "My thoughts/activities run fluidly and smoothly") and 'absorption' (e.g., "I am completely lost in

Table 5 Intercorrelations, reliability estimates (Cronbach's α and McDonald's ω), and descriptive statistics for global flow and all subscales among full dataset ($N = 913$)

Subscale	Pearson's correlations				α	ω	M	SD	Observed range	Skew	Kurtosis
	Global flow	Absorption	Effort-less control	Intrinsic reward							
Global flow	-				0.85	0.82	5.58	0.87	1 to 7	-0.79	1.07
Absorption	0.74***	-			0.87	0.88	5.67	1.12	1 to 7	-1.06	0.99
Effort-less control	0.67***	0.24***	-		0.83	0.85	5.41	1.06	1 to 7	-0.78	0.92
Intrinsic reward	0.82***	0.44***	0.32***	-	0.93	0.93	5.66	1.32	1 to 7	-1.22	1.32

***$p < .001$

thought")—were utilised because the remaining 3 questions target perceived (outcome) importance. Responses were made on a scale ranging from 1 (*not at all*) to 7 (*very much*). Responses to all 10 items were summed to determine a composite score for further analyses. The internal consistency estimate for scores derived from this measure in this study was $\alpha = 0.85$.

Activity Details We used a single item to identify the activity the participant was referring to in their experience: "Please state the task/activity that you used to answer the above questions". A text area was made available for responses. A second single item was used to identify the duration of the activity: "Please state in minutes how long the task/activity lasted for". A small text area was made available for responses.

Self-Report Flow Entry Participants were instructed, "the questions above are designed to assess 'flow' experiences. Flow is described as 'a total engagement in which nothing else matters, actions seem to flow effortlessly—simply participating in the act feels satisfying.' Based on your experience and responses to the above items, do you think that you experienced 'flow' in your recent task/activity?" Participants were asked to respond, 'yes', 'no', or 'unsure', with the aim of providing a discrete assessment of flow entry (see discussions around flow being a discrete and continuous construct; Norsworthy et al., 2021; Peifer & Engeser, 2021a, b). If participants answered 'no' to this item, they were not asked to complete the following items.

Flow Duration Participants reported the percentage of time they felt that they were in flow within the given activity using the item, "What percentage of the time were you in flow in relation to the total time (length) of the activity?" A percentage score was designated for responses.

Flow Frequency To determine flow frequency, participants were asked, "was the flow experience a single experience or did you experience it or enter and exit it on multiple occasions within the activity?" Ordinal responses ('just once', 'on two or three occasions', 'multiple occasions') were presented.

Flow Intensity Participants reported their perceptions regarding the intensity of the flow experience by responding to the question, "Referring back to the experience of flow, how strongly did you experience flow?" A response scale was provided that ranged from 1 (*very weak*) to 7 (*very strong*). See Supplementary Material for these flow items in full.

3.4.2 Results

Activity Details Participants engaged in over 100 unique types of activities, including physical exercise, running, cooking, golf, music practice, drawing, jigsaw puzzles, email responding, gaming, studying, cycling, fitness work out, housework,

aerobics, tennis, work meeting, yoga, knitting, canoeing, researching, exams, carpentry, and work-related tasks. Activities ranged from 6 min in length to 4 h.

Correlates To assess convergent and discriminant validity and examine the relationship between PFS scores and the above assessments, descriptive statistics and a series of correlation coefficients were computed with the full dataset ($n = 913$) using IBM SPSS (version 27). See Table 6 for detailed statistics. As expected, perceived anxiety and stress scores significantly negatively correlated with the global PFS score and absorption, effort-less control and intrinsic reward PFS subscale scores. Scores for perceived competence, self-reported performance, autotelic personality, flow (as measured by the FSS), flow intensity, and flow duration all significantly positively correlated with global PFS and all PFS subscale scores. Broadly, these correlation values demonstrate that the nomological network associated with PFS scores 'operates' in a way that is consistent with existing flow research, and provide preliminary evidence for the external aspect of validity for PFS scores. Correlations between the above variables and between global PFS and the subcomponents of autotelic personality are reported in Tables S5 and S6, respectively.

The examination of general flow appraisals revealed that 71% ($n = 648$) of participants reported being in flow (i.e., *flow entry*), 18% ($n = 164$) reported not being in flow, and 11% ($n = 101$) reported being unsure as to whether they were in flow or not (see Table S7 for further details). On average, participants reported being in flow 66.49% ($SD = 24.48$) of the time during their nominated activity (*flow duration*). They also reported a mean flow intensity of 6.5 out of 10 ($SD = 1.6$). See Figs. S6 and S7 for histograms of flow duration and intensity. Flow duration and flow intensity both correlated positively and significantly with global PFS scores, and the PFS subscale scores for absorption, effort-less control, and intrinsic reward. See Table S8 for information on frequency of flow within a single event. For the interested reader, a positive and significant correlation was evident between the percentage of time participants self-reported to be in flow and self-reported intensity of flow ($r = .47$, $p < .001$). Taken together, these correlations with global and subscale PFS scores—derived using simple, naturalistic assessments of flow characteristics—offer evidence that the PFS (and its dimensions) may provide insight into important elements of one's overall flow experience in a given activity.

4 Discussion

In response to recent critiques of the flow literature (e.g., Peifer & Engeser, 2021a, b; Swann et al., 2018), Norsworthy et al. (2021) conducted a scoping review and identified three core 'flow experience' themes (i.e., absorption, effort-less control, and intrinsic reward) that appear to exist across scientific disciplines and activity domains. Although the area is replete with measurement instruments— the existing

psychological assessments showed substantiative issues such as confounding ante-cedents into the experience measure, and not directly assessing the *effortlessnes*s associated with high degrees of control that specifically delineates (and not con-flates) flow from other similar states—no instrument exists to adequately assesses those three dimensions of the flow experience. The purpose of this study was to develop, and provide preliminary evidence of reliability and validity for, a new psy-chological flow scale (PFS). Overall, our findings offered insight into Messick's (1995) content, substantive, structural, and external aspects of validity regarding the PFS across a variety of activity domains, and support for further use of the PFS as a psychological assessment of flow in English speaking adults.

4.1 Advancing the Measurement of Flow

The PFS advances our measurement of flow by targeting the core *experiential* dimensions of the construct, rather than conflating flow antecedents, pre-conditions, or outcomes (which are included in many existing measures) with the experience of flow itself. Further, the PFS is intended to be used in a domain-general sense and as a result does not rely on domain-specific (e.g., telepresence in computing) descrip-tions that may not generalise across scientific disciplines, domains or nuanced experiential descriptions. We also presented evidence to show that flow may be operationalised using the PFS with a higher-order latent global factor alongside the dimensions: *absorption, effort-less control*, and *intrinsic reward*. These results may lend support to the notion that an attempt to model the flow experience *as a whole*—alongside or perhaps, in some cases, instead of the three sub-dimensions—may be valuable for informing our understanding of the construct and how it operates (see Norsworthy et al., 2021). The EFA revealed a small fourth factor involving items that best represented intuitive behaviour or spontaneous action, which seemed to be inter-related with the effort-less control dimension. Although the CFA showed evidence against its inclusion in the final PFS, future research may want to exam-ine how this 'intuiting' construct relates to effortless-control, especially within non-movement associated activities (i.e., brainstorming) in which cognitive aspects of the experience may be more prevalent. Finally, in outlining the instructions for the PFS, we attempted to bring some clarity to flow measurement by asking participants to report on the most intense experience in the given event—reducing interpretive ambiguity as to whether responses to flow items represent a conflated amalgamation (averaged or summed) of multiple experiences into a single reported experience (see Moneta, 2012; Swann et al., 2018).

4.2 External Validity

Findings revealed preliminary evidence for the nomological network associated with PFS scores, and in doing so revealed support for the convergent and discri-minant validity (Messick, 1995). More specifically, global and dimensional PFS scores were negatively correlated with perceived anxiety and stress, and positively

Table 6 Correlations between global flow and subscales and theoretically relevant variables

Variable	N	M	SD	Observed range	Pearson's correlation coefficients			
					Global flow	Absorption	Effort-less control	Intrinsic reward
Perceived anxiety	913	2.49	1.92	1 to 10	−0.33***	−0.07*	−0.36***	−0.29***
Perceived competence	913	33.94	6.21	7 to 42	0.47***	0.22***	0.56***	0.30***
Perceived stress	913	12.34	6.04	5 to 35	−0.40***	−0.11***	−0.46***	−0.33***
Self-reported performance	913	5.47	1.09	1 to 7	0.47***	0.32***	0.47***	0.29***
Autotelic personality	913	124.95	16.25	68 to 175	0.38***	0.30***	0.33***	0.24***
Flow short scale	913	52.09	9.52	12 to 70	0.73***	0.57***	0.59***	0.50***
Flow intensity	743	6.5	1.6	1 to 10	0.51***	0.39***	0.37***	0.35***
Flow duration	742	66.49	24.48	2 to 100	0.35***	0.28***	0.27***	0.22***

*** $p < .001$

correlated with perceived competence, self-rated performance, and autotelic personality. It is important to caution against making any causal inferences regarding these associations; nonetheless, they are consistent (in a directional sense) with theory and research that identify flow as a buffer to anxiety and stress (also see Waples & Knight, 2017; Llorens & Salanova, 2017; Peifer et al., 2014; & Sadlo, 2016), and facilitator of positive development and high functioning (also see Flett, 2015; Klasen et al., 2012; Norsworthy et al., 2017; Norsworthy et al., 2021, & Swann et al., 2017).

One criticism of existing flow assessments (e.g., using the FSS-2 and FSS) is that, despite there being no 'threshold' explicitly recommended by those who developed the instruments, flow is often assumed to be 'present' when scores are above a midpoint (see Jackman et al., 2017). Participants in this study who reported that they *had* experienced flow (reporting 'Yes' from a 'Yes', 'No' or 'Unsure' response option) scored higher on global PFS scores ($M = 5.86$; out of 7) than those who reported they were unsure ($M = 5.21$) or had not experienced flow ($M = 4.68$); all of these mean flow scores exceeded the midpoint (3.5). On the assumption that flow has an entry point (differentiating it from other immersive or automatic states) and also varies by intensity (accounting for mild and intense flow experiences; see Norsworthy et al., 2021), it therefore appears important to caution against 'assuming' flow occurs when scores reach or exceed a scale midpoint.

On the topic of flow 'intensity', it is worth highlighting that in this study self-reported flow intensity scores significantly positively correlated with PFS scores. As such, it is reasonable to conclude that PFS scores may align with self-reported intensities of simple descriptions of flow, and that higher PFS scores may indicate higher self-reported flow intensity. Additionally, flow intensity also correlated positively with the percentage of time that participants self-reported to be in flow, and their global PFS scores, indicating that higher intensities of flow may take time to develop (i.e., a flow experience may progressively grow in intensity, rather than instantly occur at a 'deep' level). This proposition would be consistent with the hypothesis that flow occurs through a depleting 'onion-peeling' effect of higher cognitive processes, which results in progressive absorption and greater reduction in felt effort (see Norsworthy et al., 2021; Sadlo, 2016; Ulrich et al., 2016).

4.3 Limitations and Future Directions

In this investigation, we tested the PFS with adults who were resident in multiple English speaking countries and who performed an array of focal activities. It is necessary to highlight, though, that instrument validation is an iterative process and that our results provide only preliminary insight into the development and use of the PFS in this field. With that in mind, we encourage researchers to expand the evidence base for or against the use of the PFS by examining validity and reliability properties in different cultures and languages, in adult and youth samples, within and between focal activities, and on multiple occasions within individuals. We derived insight using a cross-sectional approach; longitudinal studies, therefore, could be used to evaluate within-person variability or stability in PFS scores (and relations with

correlate variables). In addition, although our purpose was to develop a *psychological* flow scale, it would also be worthwhile—in light of emerging psychophysiological work in this area—for researchers to examine key (measurable) physiological or neuroscientific markers of flow alongside insight derived through instruments such as the PFS (see, for example, Norsworthy et al., 2021; Peifer & Tan, 2021).

4.4 Conclusion

The Psychological Flow Scale (PFS) is a relatively brief instrument that assesses three related dimensions (i.e., absorption, effortless-control, intrinsic reward) that are proposed to characterise the flow state experience across different domains and contexts. This instrument is an important advancement for the field of flow science because (a) it may enable valuable cross-disciplinary comparisons to be made in future work, (b) it is based upon the findings of a recent synthesis that aimed to provide a parsimonious understanding of the construct, (c) it removes what researchers believe to be the 'antecedents' (or operational factors) from assessing the experience of flow, and (d) addresses issues of construct conflation (although further discriminant validity testing is required) with similar states. In addition, evidence derived from CFA offers preliminary support for a higher-order flow construct that addresses recent conceptual and measurement criticisms of flow. We encourage sustained efforts focused on developing the PFS and demonstrating its applicability and correlates across populations and activities. As well as advancing our assessment and understanding of people's flow experiences, we also hope the PFS will offer practical value for researchers and intervention designers in terms of quantifying our efforts to improve people's experiences (and to 'find flow') in diverse activities.

Appendix

Psychological Flow Scale (PFS)

The below questions relate to the thoughts and feelings you may have experienced while taking part in your recent activity. There are no right or wrong answers. Think about how you felt during the event/activity, then answer the questions using the rating scale.

Please rate the questions below in relation to the most intense optimal moment you experienced in your given event.

		Strongly disagree	Neutral					Strongly agree
1	I was absorbed in the act/task	1	2	3	4	5	6	7
2	I was highly focused on the task/activity	1	2	3	4	5	6	7
3	All my attention was on the task/activity	1	2	3	4	5	6	7
4	I felt like I could easily control what I was doing	1	2	3	4	5	6	7
5	My actions flowed effortlessly	1	2	3	4	5	6	7

		Strongly disagree	Neutral						Strongly agree
6	There was a sense of fluidity to my actions	1	2	3	4	5	6	7	
7	I found the experience rewarding	1	2	3	4	5	6	7	
8	The experience felt satisfying	1	2	3	4	5	6	7	
9	I would like the feeling of that experience again	1	2	3	4	5	6	7	

Researcher notes:

Items 1-9 assess the flow experience. Average the nine scores to obtain a global flow score. Items 1-3 assess the dimension 'absorption'. Items 4-6 assess the dimension 'effort-less control'. Items 7-9 assess the dimension 'intrinsic reward'. Where possible it is advised to examine dimensional scores in addition to the global flow score. Avg means scores (for both global and the three dimensions) to be used for reporting.

Declarations

This study adhered to all ethical standards.

Conflict of interest The authors report no conflict of interests. The funding source played no further role in the study design, the collection, analysis, and interpretation of data and in writing report of this work.

References

Abuhamdeh, S. (2021). On the relationship between flow and enjoyment. In S. Engeser, & C. Peifer (Eds.), *Advances in flow research* (pp. AQ: 18 155–169). Springer. https://doi.org/10.1007/978-3-030-53468-4_6

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: review and recommendations. *Journal of Management, 47*(4), 823–837.

Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods, 24*(4), 468.

Auld, D. P. (2014). *Flow and learning in computer-mediated learning environments. A Meta-Analytic Review* [Doctoral dissertation]. Fordham University. https://doi.org/10.1037/e549562014-001

Barthelmäs, M., & Keller, J. (2021). Antecedents, boundary conditions and consequences of flow. In S. Engeser, & C. Peifer (Eds.), *Advances in flow research* (pp. 71–107). Springer. https://doi.org/10.1007/978-3-030-53468-4_3

Baumann, N. (2012). Autotelic personality. In S. Engeser (Ed.), *Advances in flow research* (pp. 165–186). Springer. https://doi.org/10.1007/978-1-4614-2359-1_9

Bian, Y., Yang, C., Gao, F., Li, H., Zhou, S., Li, H., Sun, X., & Meng, X. (2016). A framework for physiological indicators of flow in VR games: construction and preliminary evaluation. *Personal and Ubiquitous Computing, 20*(5), 821–832. https://doi.org/10.1007/s00779-016-0953-5

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science, 13*(2), 149–154.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. Jossey-Bass. https://doi.org/10.1037/h0098761.

Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi*. Springer. https://doi.org/10.1007/978-94-017-9088-8

Csikszentmihalyi, M., Montijo, M. N., & Mouton, A. R. (2018). Flow theory: optimizing elite performance in the creative realm. In S. I. Pfeiffer, E. Shaunessy-Dedrick, & M. Foley-Nicpon (Eds.), *APA handbooks in psychology®. APA handbook of giftedness and talent* (pp. 215–229). American Psychological Association. https://doi.org/10.1037/0000038-014.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research, 50*(1), 61–104.

de Manzano, Ö, Cervenka, S., Jucaite, A., Hellenäs, O., Farde, L., & Ullén, F. (2013). Individual differences in the proneness to have flow experiences are linked to dopamine D2-receptor availability in the dorsal striatum. *NeuroImage, 67*(1), 1–6.

Dietrich, A. (2004). Neurocognitive mechanisms underlying the experience of flow. *Consciousness and Cognition, 13*(4), 746–761. https://doi.org/10.1016/j.neuroimage.2012.10.072

Dunn, K. J., & McCray, G. (2020). The place of the bifactor model in confirmatory factor analysis investigations into construct dimensionality in language testing. *Frontiers in Psychology, 11*, 1357. https://doi.org/10.3389/fpsyg.2020.01357

Ellis, G. D., Freeman, P. A., Jiang, J., & Lacanienta, A. (2019). Measurement of deep structured experiences as a binary phenomenon. *Annals of Leisure Research, 22*(1), 119–126. https://doi.org/10.1080/11745398.2018.1429285

Engeser, S., & Schiepe-Tiska, A. (2012). Historical lines and an overview of current research on flow. In S. Engeser (Ed.), *Advances in flow research* (pp. 1–22). Springer. https://doi.org/10.1007/978-1-4614-2359-1_1

Engeser, S. E. (2012). Theoretical integration and future lines of flow research. In S. Engeser (Ed.), *Advances in flow research* (pp. 187–199). Springer. https://doi.org/10.1007/978-1-4614-2359-1_10

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Flett, M. R. (2015). Is flow related to positive feelings or optimal performance? Path analysis of challenge-skill balance and feelings. *Sport Science Review, 24*(1–2), 5–26. https://doi.org/10.1515/ssr-2015-0006

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. https://doi.org/10.1080/17439760.2014.967799

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59*(2), 93. https://doi.org/10.1037/0003-066x.59.2.93

Harris, D. J., Vine, S. J., & Wilson, M. R. (2017). Neurocognitive mechanisms of the flow state. *Progress in Brain Research, 234*, 221–243. https://doi.org/10.1016/bs.pbr.2017.06.012

Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: a functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238. https://doi.org/10.1037/1040-3590.7.3.238

Heutte, J., Fenouillet, F., Kaplan, J., Martin-Krumm, C., & Bachelet, R. (2016). The EduFlow model: A contribution toward the study of optimal learning environments. In L. Harmat, F. Ø. Andersen, F.

Ullén, J. Wright, & G. Sadlo (Eds.), *Flow experience* (pp. 127–143). Springer. https://doi.org/10.1007/978-3-319-28634-1_9

Heutte, J., Fenouillet, F., Martin-Krumm, C., Gute, G., Raes, A., Gute, D., Bachelet, R., & Csikszent-mihalyi, M. (2021). Optimal experience in adult learning: conception and validation of the Flow in Education Scale (EduFlow-2). *Frontiers in Psychology, 12*, 1–12.

Hoyle, R. H. (2011). *Structural equation modeling for social and personality psychology*. Sage.

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 and Exercise, 32*(1), 113–123. https://doi.org/10.1016/j.psychsport.2017.06.009

Jackson, S. A., & Csikszentmihalyi, M. (1999). *Flow in sports*. Human Kinetics.

Jackson, S. A., & Eklund, R. C. (2002). Assessing flow in physical activity: The flow state scale-2 and dispositional flow scale-2. *Journal of Sport & Exercise Psychology, 24*(2), 133–150. https://doi.org/10.1123/jsep.24.2.133

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36.

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement, 34*(1), 111–117.

Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., & Mathiak, K. (2012). Neural contributions to flow experience during video game playing. *Social Cognitive and Affective Neuroscience, 7*(4), 485–495. https://doi.org/10.1093/scan/nsr021

Krane, V. (1994). The mental readiness form as a measure of competitive state anxiety. *The Sport Psychologist, 8*(2), 189–202. https://doi.org/10.1123/tsp.8.2.189

Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 9*(08), 2207. https://doi.org/10.4236/psych.2018.98126

Lazoc, A., & Luiza, C. (2012). Elaborating a measurement instrument for the flow experience during online information search. *Annals of the University of Oradea Economic Science Series, 21*(2), 841–847.

Llorens, S., & Salanova, M. (2017). The consequences of flow. In A. B. Bakker, & M. van Woerkom (Eds.), *Flow at work: Measurement and implications*. Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315871585-6

Llorens, S., Salanova, M., & Rodríguez, A. M. (2013). How is flow experienced and by whom? Testing flow among occupations. *Stress and Health, 29*(2), 125–137. https://doi.org/10.1002/smi.2436

Martens, R., Vealey, R. S., Burton, D., Bump, L., & Smith, D. E. (1990). Development and validation of the Competitive Sports Anxiety Inventory 2. In R. Martens, R. S. Measures of anxiety 16 Vealey, & D. Burton (Eds.), *Competitive anxiety in sport* (pp. 117–178). Human Kinetics. https://doi.org/10.1037/t27557-000

Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Moneta, G. B. (2012). On the measurement and conceptualization of flow. In S. Engeser (Ed.), *Advances in flow research* (pp. 23–50). Springer. https://doi.org/10.1007/978-1-4614-2359-1_2

Moran, A. (2012). Concentration: attention and performance. In S. Murphy (Ed.), *The Oxford handbook of sport and performance psychology* (pp. 117–130). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199731763.013.0006.

Nah, F. F. H., Yelamanchili, T., & Siau, K. (2017). A review on neuropsychophysiological correlates of flow. In F. F. H. Nah, & C. H. Tan (Eds.), *International conference on HCI in business, government, and organizations* (pp. 364–372). Springer.

Norsworthy, C., Thelwell, R., Weston, N., & Jackson, S. A. (2017). Flow training, flow states, and performance in élite athletes. *International Journal of Sport Psychology, 49*(1), 134–152. https://doi.org/10.7352/IJSP.2018.49.134

Norsworthy, C., Jackson, B., & Dimmock, J. A. (2021). Advancing our understanding of psychological flow: a scoping review of conceptualizations, measurements, and applications. *Psychological Bulletin, 147*(8), 806–827. https://doi.org/10.1037/bul0000337

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., & Yarkoni, T. S. (2015). C. I. E. N. T. I. F. I. C. S. T. A. N. D. A. R. D. S. Promoting an open research culture. *Science*, *348*(6242), 1422–1425.

O'Connor, B. P. (2023). *EFA.dimensions: Exploratory factor analysis functions for assessing dimensionality* (version 0.1.7.6) [R package]. https://CRAN.R-project.org/package=EFA.dimensions

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22–27.

Peifer, C., Schulz, A., Schächinger, H., Baumann, N., & Antoni, C. H. (2014). The relation of flow-experience and physiological arousal under stress—can u shape it? *Journal of Experimental Social Psychology, 53*(1), 62–69. https://doi.org/10.1016/j.jesp.2014.01.009

Peifer, C. (2012). Psychophysiological correlates of flow-experience. In S. Engeser (Ed.), *Advances in flow research* (pp. 139–164). Springer. https://doi.org/10.1007/978-1-4614-2359-1_8

Peifer, C., & Engeser, S. (2021a). *Advances in flow research*. Springer. https://doi.org/10.1007/978-3-030-53468-4

Peifer, C., & Engeser, S. (2021b). Theoretical Integration and future lines of flow research. In S. Engeser & C. Peifer (Eds.), *Advances in flow research* (pp. 417–439). Springer. https://doi.org/10.1007/978-3-030-53468-4_16

Peifer, C., & Tan, J. (2021). The psychophysiology of flow experience. In S. Engeser & C. Peifer (Eds.), *Advances in flow research* (pp. 191–230). Springer. https://doi.org/10.1007/978-3-030-53468-4_8

Peifer, C., Wolters, G., & Harmat, L. (2022). A scoping review of flow research. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2022.815665

Rheinberg, F., Engeser, S., & Vollmeyer, R. (2002). Measuring components of flow: the flow-short-scale. *Proceedings of the 1st International Positive Psychology Summit*. https://doi.org/10.1037/t47787-000

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*(3), 450. https://doi.org/10.1037/0022-3514.43.3.450

Romero, P., & Calvillo-Gamez, E. (2014). An embodied view of flow. *Interacting with Computers, 26*(6), 513–527. https://doi.org/10.1093/iwc/iwt051

Sadlo, G. (2016). Towards a neurobiological understanding of reduced self-awareness during flow: An occupational science perspective. In L. Harmat, F. Ø. Andersen, F. Ullén, J. Wright, & G. Sadlo (Eds.), *Flow Experience* (pp.375–388). Springer. https://doi.org/10.1007/978-3-319-28634-1_22

Schüler, J., & Brandstätter, V. (2013). How basic need satisfaction and dispositional motives interact in predicting flow experience in sport. *Journal of Applied Social Psychology, 43*(4), 687–705. https://doi.org/10.1111/j.1559-1816.2013.01045.x

Shin, N. (2006). Online learner's 'flow' experience: An empirical study. *British Journal of Educational Technology, 37*(5), 705–720. https://doi.org/10.1111/j.1467-8535.2006.00641.x

Sinnamon, S., Moran, A., & O'Connell, M. (2012). Flow among musicians: measuring peak experiences of student performers. *Journal of Research in Music Education, 60*(1), 6–25. https://doi.org/10.1177/0022429411434931

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. https://doi.org/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. https://doi.org/10.1016/j.psychsport.2015.10.008

Swann, C., Crust, L., & Vella, S. A. (2017). New directions in the psychology of optimal performance in sport: Flow and clutch states. *Current Opinion in Psychology, 16*(1), 48–53. https://doi.org/10.1016/j.copsyc.2017.03.032

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

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Tozman, T., Magdas, E. S., MacDougall, H. G., & Vollmeyer, R. (2015). Understanding the psychophysiology of flow: a driving simulator experiment to investigate the relationship between flow and heart rate variability. *Computers in Human Behavior, 52*(1), 408–418. https://doi.org/10.1016/j.chb.2015.06.023

Tse, D. C., Lau, V. W. Y., Perlman, R., & McLaughlin, M. (2020). The development and validation of the Autotelic Personality Questionnaire. *Journal of Personality Assessment, 102*(1), 88–101. https://doi.org/10.1080/00223891.2018.1491855

Tse, D. C., Nakamura, J., & Csikszentmihalyi, M. (2021). Living well by "flowing' well: the indirect effect of autotelic personality on well-being through flow experience. *The Journal of Positive Psychology, 16*(3), 310–321. https://doi.org/10.1080/17439760.2020.1716055

Ulrich, M., Keller, J., & Grön, G. (2016). Dorsal raphe nucleus downregulates medial prefrontal cortex during experience of flow. *Frontiers in Behavioral Neuroscience, 10*(1), 169. https://doi.org/10.3389/fnbeh.2016.00169

Valenzuela, R., Codina, N., & Pestana, J. V. (2018). Self-determination theory applied to flow in conservatoire music practice: the roles of perceived autonomy and competence, and autonomous and controlled motivation. *Psychology of Music, 46*(1), 33–48. https://doi.org/10.1177/0305735617694502

Waples, C., & Knight, P. (2017). *Flow in the context of industrial and organizational psychology: the case of work motivation In Flow at work* (pp. 152–168). Routledge. https://doi.org/10.4324/9781315871585-8

Zito, M., Cortese, C. G., & Colombo, L. (2018). The italian adaptation of the WOrk-reLated Flow Inventory (WOLF) to Sport: the I-WOLFS scale. *BPA-Applied Psychology Bulletin, 66*(281), 38–45. https://doi.org/10.1037/t68092-000

Authors and Affiliations

Cameron Norsworthy[1] [ORCID] **· James A. Dimmock[2] · Daniel J. Miller[2] · Amanda Krause[2] · Ben Jackson[1,3]**

✉ Cameron Norsworthy
cameron.norsworthy@research.uwa.edu.au; cameron@flowcentre.org

James A. Dimmock
james.dimmock@jcu.edu.au

Daniel J. Miller
daniel.miller1@jcu.edu.au

Amanda Krause
amanda.krause1@jcu.edu.au

Ben Jackson
ben.jackson@telethonkids.org.au

[1] School of Human Sciences, The University of Western Australia, Perth, Australia

[2] Department of Psychology, College of Healthcare Sciences, James Cook University, Townsville, Australia

[3] Telethon Kids Institute, Perth, Australia