

# How Far Does Flow Get You? Disentangling the Within-Person, Between-Person, and Directional Relationship Between Flow, Task Difficulty, and Performance

Birte Ann-Kathrin Thissen

bt2357@nyu.edu

New York University

Mark Lustig

Universität Hamburg

Bumin Ko

New York University

Moritz Burghardt

Heidelberg University

Gabriele Oettingen

New York University

---

## Research Article

**Keywords:** flow state, objective performance, subjective performance, task difficulty, multilevel modeling

**Posted Date:** October 14th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-11059958/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

## Abstract

Flow is defined as an intrinsically rewarding state of complete task engagement and is thus often assumed to also enhance task performance. To examine the nature of the flow-performance relationship, we conducted a controlled lab study in which students ( $N = 257$ ) played two sessions of the computer game *Tetris*. Participants were randomized into a balance condition with matching task difficulty or an overload condition with high task difficulty. After each session, they reported their flow and gave a rating of how well they had played (subjective performance), while their point score was also being recorded (objective performance). Multilevel models showed that flow positively predicted both subjective and objective performance beyond practice effects, task difficulty, and interindividual differences. Experiencing more flow than usual as well as experiencing more flow than others were each linked to better subjective and objective performance. Multilevel mediation models suggested that flow may be one pathway through which task difficulty affects performance, as flow partially mediated its effect on objective performance and had an indirect effect on subjective performance. However, we found no evidence for a directional relationship, as flow in the first session did not predict performance in the second session when controlling for baseline and task difficulty, nor vice versa. Across analyses, flow was more strongly associated with subjective than with objective performance. Together, these findings suggest that flow and performance are positively related during immediate task execution, and that this relationship may reflect a coupling of shared task engagement processes.

## Introduction

Ever since Mihaly Csikszentmihalyi proposed the concept of *flow* in 1975, motivational psychologists have been interested in this state not only due to its close link to intrinsic motivation, but also due to its potential positive implications for human functioning. Since flow is defined as an inherently rewarding state of complete task engagement, characterized by heightened attentional focus, seemingly effortless task processing, and inherent enjoyment (Csikszentmihalyi, 1975; Jackson & Eklund, 2002), it bears an evident theoretical connection to enhanced task performance: On the one hand, being in a state of flow could facilitate momentary performance of a given task by keeping the person highly focused, engaged, and motivated; on the other hand, flow could also become a catalyst for the person to become better at their task over time through sustained motivation and positive carry-over effects (Barthelmäs & Keller, 2021). As flow is thought to arise primarily under task difficulty conditions that are optimally balanced with the person's skill level (Csikszentmihalyi, 1975; Fong et al., 2015), flow is also, by definition, linked to situations that are challenging, but not too challenging—in other words, situations that can motivate excellent performance by providing the person with a sense of mastery (Thissen & Oettingen, 2026).

While these theoretical considerations strongly suggest a positive relationship between flow and performance, empirical results are more mixed across different tasks (Cseh et al., 2015; Schüller & Brunner, 2009), implying that this relationship is more nuanced than a simple equation of flow with enhanced task performance. Even though recent meta-analyses in the contexts of sports and gaming (Harris et al., 2023) and learning (Jinmin & Qi, 2023) concluded that flow and performance are moderately positively related, the exact nature of this relationship—its underlying mechanisms, causal direction, and conceptualization—remains largely unclear (Harris et al., 2023). Harris and colleagues (2023) point out two major methodological challenges that complicate the empirical study of the flow-performance relationship and limit our current insights: First, with no clear objective online marker for flow in place, flow assessment typically relies on retrospective self-reports, which could be positively biased after an excellent performance; therefore, it is important to distinguish between associations of flow with objective and with subjective performance. Second, as most studies to date do not control for other performance-relevant factors, the reported relationship between flow and performance could be obfuscated or confounded by third variables. For instance, without taking into account interindividual skill differences and examining both the within-person and between-person level, it cannot be inferred how robust the positive relationship between flow and performance really is, or whether it is simply driven by highly skilled performers also being more prone to flow rather than by a more general association. This second point also speaks to the larger methodological issue of having to rely, for the most part, on correlational research, as the established manipulation approach for flow experiments—varying the task's difficulty level to compare flow-conducive balanced to flow-adverse imbalanced conditions—affects not only flow but also performance, potentially confounding results (Barthelmäs & Keller, 2021). Thus, causality and experimental control over third variables are difficult to establish in this line of research, and what at first glance seems like a straightforward relationship becomes rather challenging to disentangle and truly understand.

Against this background, the current study was aimed at illuminating the nature of the flow-performance relationship, by testing its robustness with regard to other performance-relevant factors like task difficulty, practice effects, and interindividual differences, and by decomposing within-person and between-person level associations. Moreover, a theoretically plausible underlying mechanism was tested, namely that flow serves as a mediator for task difficulty effects on performance. Finally, directional relationships between flow

and performance were tested, to explore short-term carry-over effects. It should be noted that these are secondary analyses of an existing dataset from a lab study that also tested the effect of mood induction on flow; however, since the mood induction failed to produce the intended effect on mood and did not affect flow or performance (see Supplement S1 to S3), mood induction groups were pooled for these secondary analyses. To ensure the validity of this approach, participants' mood group was entered as an additional predictor variable into sensitivity analyses for each model.

## Flow and Task Performance

The flow concept goes back to highly intrinsically motivated individuals describing what engagement with "their" activity feels like, pointing out several experiential phenomena that accumulate in a distinct, holistic state of mind: perceived skill-challenge balance, a sense of control or competence, an unambiguous understanding of the task, clear goals, heightened attentional focus, reduced self-awareness, an altered sense of time, a merging between action and awareness, and intrinsic enjoyment of the task at hand (Csikszentmihalyi, 1975; Jackson & Eklund, 2002). While there is an ongoing scientific debate about which of these components form the core of the flow experience (Norsworthy et al., 2023), it seems straightforward that most, if not all, can be theoretically linked to task performance.

On the one hand, the attention-related aspects of flow imply better attentional performance, which first psychophysiological (Harris et al., 2017) and neuropsychological findings (de Sampaio Barros et al., 2018) as well as studies involving validated attention tests (Marty-Dugas et al., 2021; Thissen & Oettingen, 2026) corroborate. On the other hand, the processing-related aspects of flow evoke a sense of effortlessness, which combined with intrinsic task enjoyment, illustrates the close link between flow and intrinsic motivation. Indeed, a recent meta-analysis on flow and intrinsic motivation in students (Mehta & Vyas, 2022) reports a strong concurrent relationship, and task-related intrinsic motivation has been found to increase after a flow experience (Keller et al., 2011; Schüler & Brunner, 2009). Moreover, a meta-analysis by Fong and colleagues (2015) demonstrated that balanced task difficulty, besides offering favorable conditions to apply one's skills efficiently, relates positively to both flow and intrinsic motivation, with the former showing the stronger relationship. Considering that intrinsic motivation is a medium-to-strong predictor of performance according to a meta-analysis of over four decades of research (Cerasoli et al., 2014), these findings suggest that flow states are also linked to performance enhancements through motivational benefits. For instance, Thissen and Oettingen (2026) found that effort exertion is perceived in a more positive manner in conjunction with flow experiences, as they provide a sense of mastery over challenge.

Despite flow showing evident connections to task performance through its very conceptualization, a positive flow-performance relationship has not been found for every task domain: For instance, null-associations between flow and objective performance have been reported for prominent flow activities such as creative tasks (Cseh et al., 2015; Gaggioli et al., 2017), running (Schüler & Brunner, 2009; Stoll & Lau, 2005), and online gaming (Schiefele & Roussakis, 2006), whereas the association between flow and subjective performance seems to be more consistent (Cseh et al., 2015; Thissen & Oettingen, 2026). Some evidence also suggests that the apparent association between flow and performance could be explained by other factors facilitating both flow and performance at the same time, as controlling for task difficulty conditions resulted in null associations in flow experiments (Keller & Bless, 2008; Keller & Blomann, 2008). Moreover, while some studies focus on the situational and momentary association between flow and performance (Aitken et al., 2023; Fullagar & Kelloway, 2009; Stollberger & Debus, 2020), others emphasize stable differences across individuals as the basis of the flow-performance relationship (Engeser & Rheinberg, 2008; Schüler & Brunner, 2009). Yet, to the best of our knowledge, no study to date has decomposed the relationship into within- and between-person components, to gain a better understanding of how it primarily unfolds. In a similar vein, few studies to date have tested the mechanisms through which the flow-performance relationship unfolds (Harris et al., 2023). While flow has been identified as a mediator for performance effects of certain task conditions, such as environmental resources (Bakker et al., 2011) or multitasking demands (Peifer & Zipp, 2019), the flow-performance relationship itself appears to be partially mediated by prior experience with the task (Palomäki et al., 2021). However, the relatively straightforward notion that task difficulty does not only have a distinct effect on both flow and performance, but also affects one through the other, has not been tested, yet. When a positive flow-performance relationship unfolds under balanced task conditions, it seems plausible that flow, at least in part, mediates the positive task difficulty effect on performance, as, for instance, flow also mediates task difficulty effects on effort exertion (Lu et al., 2025) and performance anxiety (Fullagar et al., 2013). Perhaps the biggest open research question with regard to flow and performance to date, however, concerns causality, given the aforementioned methodological challenges to experimentally test such effects. These methodological challenges notwithstanding, exploring the directionality of the flow-performance relationship could provide important first insights. Yet, previous research in this regard has produced mixed results, reporting both positive (Christandl et al., 2018) and negative (Lavoie & Main, 2024; Liu et al., 2022) as well as reciprocal (Salanova et al., 2014) carry-over effects of flow on task performance. Therefore, further research is needed to further disentangle the flow-performance relationship.

## Current Study

The current study seeks to address the previous mixed results and open research questions described above, by examining the flow-performance relationship across two sessions of the computer game Tetris, and by employing multilevel modeling, multilevel mediation analysis, and lagged regression modeling to disentangle it. Specifically, we hypothesize that flow will positively predict both objective (H1a) and subjective performance (H1b) beyond other performance-relevant factors including task difficulty, practice effects, and interindividual differences, and that these positive associations on both the objective (H2a) and subjective (H2b) level can be observed within- as well as between-persons. Moreover, we surmise that flow serves as a mediator for the positive effect of balanced task difficulty on objective (H3a) and subjective (H3b) performance, and that flow has a directional positive relationship on objective (H4a) and subjective (H4b) performance across sessions.

## Methods

### Participants

Participants were recruited using the local university's participant pool and completed the study in exchange for course credit. Eligibility criteria included fluency in the language the survey was in, having no impairments in the dominant hand that could obstruct playing Tetris, and having no autism or affective disorder diagnosis that could affect the mood induction (not the focus of the current article). The initial sample size was  $N = 310$ ; however, 19 participants were screened out for failing the eligibility criteria. Moreover, 34 participants were excluded prior to the analyses: 23 due to failed attention checks, four due to withdrawal of consent after the debriefing, and seven due to technical issues. Thus, the final sample size was  $N = 257$ . Even though sample size considerations had been based on power estimations for testing the mood induction effect, a power analysis for the associations between flow and performance was later conducted using the *simr*-package (Green & McLeod, 2016) in the statistics software R (version 4.5.1; R Core Team, 2025). Specifically, 1,000 datasets were simulated for the fitted multilevel models and sensitivity to detect even small effects was evaluated by reducing the estimated coefficient for the association between flow and performance to 60% of its observed value (i.e., from  $b = 321.27$  to  $b = 192.76$ ) and estimating the power for detecting this fixed effect at  $\alpha = .05$ . This simulation indicated that the final sample size provided 80% power to detect an effect of this magnitude, which corresponds to an effect size of approximately  $\eta^2 = .015$ .

Most participants identified as female ( $n = 175$ ; 68%), in addition to 78 participants (30%) identifying as male, and four participants (2%) identifying as non-binary or other. Age ranged from 18 to 23 years ( $M = 19.6$ ;  $SD = 1.1$ ). When asked to select all racial categories they fell into, more than half of the sample ( $n = 147$ ; 57%) selected Asian, followed by White ( $n = 76$ ; 30%) and Black ( $n = 29$ ; 11%), with another 17 participants selecting multiracial, one participant selecting Native American, and 16 participants stating to not fall into any of these racial categories. Moreover, 16% of the sample identified as ethnically Latinx. While 17 participants had already obtained a degree, 50% were undergraduate students in their first year, and the remaining 43% were more advanced undergraduate students. Most participants ( $n = 171$ ; 67%) reported having prior familiarity with Tetris and rated their enjoyment of the game as rather high ( $Mdn = 5$ ) but their skill level as modest ( $Mdn = 3$ ) on 7-point Likert-type response scales.

### Design and Procedure

Ethics approval was obtained by the local university's institutional review board (IRB-FY2024-9087). The study was preregistered on AsPredicted under #190381, whereas the secondary analyses this article is focused on were later preregistered on the Open Science Framework under [https://osf.io/67qjg/overview?view\\_only=6ea3011e46d34fcaadc94686b861183](https://osf.io/67qjg/overview?view_only=6ea3011e46d34fcaadc94686b861183). The study had a mixed factorial  $2 \times 2 \times 2$  design, with mood induction (positive vs. neutral) and task difficulty (balance vs. overload) as independent between-subjects variables and flow as well as subjective and objective performance while playing Tetris as dependent variables, measured repeatedly (pre- vs. post-mood induction) in both task difficulty conditions.

After giving informed consent, participants answered eligibility screening questions, and, in case they were eligible, proceeded to answer demographic questions. Then, they received instructions on how to play Tetris, accompanied by images of the game and the keys that needed to be pressed, and played a two-minute practice session at the balanced difficulty level. Afterward, participants rated their skill level and enjoyment of Tetris and indicated if they had already been familiar with this game. Before starting the first six-minute Tetris session, participants were told that task difficulty could differ from the practice session and were randomized into either the balance or overload condition. Once the session was over, they reported their flow experiences, perceived skill-challenge level, and subjective performance while playing. Then, participants' baseline mood was assessed, they were randomly assigned to watch brief video clips intended to induce either positive or neutral mood, and their mood was re-assessed. The study ended with a second six-

minute Tetris session and repeated measurements for flow, perceived skill-challenge level, and subjective performance, followed by the participant debriefing.

## Measures

As the mood induction is not the focus of this article, a description of the mood induction procedure and stimuli can be found in Supplement S1.

## Background

Single items were used to assess participants' age, gender, race and ethnicity, and educational background. Moreover, after having practiced playing Tetris for two minutes, participants rated how much they generally enjoy Tetris on a 7-point Likert-type response scale ranging from *not very much* to *very much* and how skilled they think they are at Tetris on a 7-point Likert-type response scale from *very bad* to *very good*. Moreover, they answered a binary forced-choice item, indicating whether they had (yes) or had not (no) played Tetris before participating in this study.

## Experimental Task

The computer game Tetris (Pajitnov, 1984), played in two six-minute sessions, served as the experimental task, as it is a classic paradigm in experimental flow research (Keller & Bless, 2008; Keller & Blomann, 2008). When playing Tetris, the player has to arrange differently shaped building blocks that continuously descend into the playing field into rows by rotating and shifting them. The game is over once the building blocks stack up to the upper limit of the playing field (i.e., up to 20 rows). The player can avert this by completing rows, as completed rows disappear.

Participants began playing after a five-second countdown and could rotate, shift, and speed up the descending building blocks using their keyboard's arrow keys. Their current score was continuously displayed at the upper corner of the screen next to the playing field. In case a game was over, the next one automatically started after another five-second countdown, until six minutes were up. The balance and overload conditions differed in the speed at which the building blocks descended. In both conditions, they initially dropped one row every 0.5 seconds. In the balance condition, this interval changed in increments of 0.05 seconds to match the player's performance, decreasing to a minimum 0.2 seconds when performance was good and increasing to up to 0.8 seconds when performance was poor. In the overload condition, the interval decreased from the initial 0.5 seconds to a minimum 0.1 seconds when performance was good and a maximum of 0.4 seconds when performance was poor. The JavaScript code used to implement the Tetris game is available at <https://github.com/markclustig/javascript-tetris>.

## Flow

Flow was measured using the 36-item Flow State Scale-2 (FSS; Jackson & Eklund, 2002), which assesses the nine flow components with four items each. Participants rated the items on a 5-point Likert-type response scale ranging from *strongly disagree* to *strongly agree*. For both flow measurements, internal consistency was excellent (session 1:  $\omega = .91$ ; session 2:  $\omega = .94$ ), and mean scores were calculated across all items, respectively.

## Performance

**Objective Performance.** Objective performance was gauged using a point-based scoring system similar to what is used in commercial versions of Tetris. Players received 10 points each time a row dropped by itself, in addition to receiving points for completing rows: 100 points for completing a single row at a time, 200 points for completing two rows simultaneously, 400 points for three simultaneous rows, and 800 points for four rows. The mean point score calculated across all games a participant played during a given Tetris session served as the objective performance measure in our analyses.

**Subjective Performance.** Subjective performance was measured after each of the two Tetris sessions using a single-item ("How well do you think you did in the game?") and a 7-point Likert-type response scale ranging from *not well* to *very well*.

## Manipulation Check

As a manipulation check for the task difficulty conditions, perceived skill-challenge balance was assessed with three items adapted from Rheinberg et al. (2003). The items were presented after each Tetris session together with 9-point Likert-type response scales ranging from 1 (*too low/easy*) to 9 (*too high/hard*), with the middle category (*just right*) indicating that the skill-challenge balance was perceived as optimal. After recoding each item so that the middle category yielded the highest score, mean scores were calculated across items for each session. For both measurements, internal consistency was acceptable given the scale length (session 1:  $\omega = .55$ ; session 2:  $\omega = .66$ ).

## Statistical Analyses

All statistical analyses were conducted in R (version 4.5.1; R Core Team, 2025), using the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages for multilevel modeling and the piecewiseSEM-package (Lefcheck, 2016) for multilevel mediation.

Multilevel modeling was employed to test H1a to H2b given substantial clustering in the performance data ( $ICC = .62$ ), as indicated by null-model results. First, in two separate models, subjective and objective performance were predicted by grand-mean-centered flow, session (second vs. first), and condition (balance vs. overload). Then, a second set of models was tested that differed from the first one only in that the flow predictor was decomposed into a between-person component (participants' grand-mean-centered mean flow) and a within-person component (participants' person-mean-centered flow). Predictors were entered into each model as fixed effects, in addition to random intercepts for participants (random slopes were not added, as this rendered the models unidentifiable due to having only two measurement points). This random-intercept specification implied a compound-symmetry covariance structure. Model parameters were estimated using restricted maximum likelihood (REML), and fixed effects were tested using Type III *F*-tests with Satterthwaite approximation.

Multilevel mediation analyses were conducted by fitting mixed-effects models within a piecewise structural equation modeling framework (Lefcheck, 2016) to test H3a and H3b. For each performance indicator, the outcome models predicted performance by condition, session, and flow, while the mediator model predicted flow by condition and session. Random intercepts for participants were included in all models. Modeling flow as the mediator was based on theoretical considerations, as the retrospective measurements did not reflect a temporal sequence. Indirect effects were calculated as the product of the estimated *a* and *b* paths, and given the non-normal sampling distribution, 95% confidence intervals were obtained via Monte Carlo simulation with 10,000 draws.

H4a and H4b were tested using lagged linear regression models: For both objective and subjective performance, the second session scores were predicted by baseline (i.e., first session) scores, condition, and flow in the first session. Model assumption checks based on diagnostic plots, Cook's distance, and variance inflation factors (VIFs) indicated heteroscedasticity, the presence of several influential data points, and residual non-normality. Therefore, heteroscedasticity-consistent (HC3) standard errors were used, and models were re-run excluding influential data points as well as using robust regression. As an exploratory analysis to test the reverse directionality, flow in the second session was predicted by a corresponding set of lagged regression models.

Sensitivity analyses included re-running all models without outliers (defined as data points 1.5 times the inter-quartile range below the first/above the third quartile) and adding Tetris-related background variables (enjoyment, skill level, and familiarity) as well as mood induction group as control variables.

## Results

Descriptive statistics are displayed in Table 1. Mean flow scores were above the response scale midpoint, showing that the Tetris game elicited a moderate-to-high level of flow in the average participant. Mean subjective performance scores, on the other hand, were below the response scale midpoint, suggesting that participants tended to rate their own performance rather conservatively. Objective performance scores varied widely across participants, ranging from 96 to 7,920, with most participants obtaining modest scores and a small number of participants obtaining very high scores (skewness = 5.12). As shown in Table 2, the first and second measurements of flow, subjective performance, and objective performance were strongly correlated, respectively, suggesting relative stability over time. Flow and subjective performance were strongly positively correlated, while flow and objective performance were moderately positively correlated.

Table 1  
*Descriptive Statistics for Flow and Performance*

| Variable                                                                                                                                                                                                                             | Total    |           | Balance condition |           | Overload condition |           |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|-----------|-------------------|-----------|--------------------|-----------|
|                                                                                                                                                                                                                                      | <i>M</i> | <i>SD</i> | <i>M</i>          | <i>SD</i> | <i>M</i>           | <i>SD</i> |
| Flow                                                                                                                                                                                                                                 | 3.28     | 0.51      | 3.34              | 0.52      | 3.21               | 0.49      |
| Objective performance                                                                                                                                                                                                                | 598      | 819       | 814               | 1,043     | 380                | 400       |
| Subjective performance                                                                                                                                                                                                               | 2.86     | 1.25      | 3.01              | 1.31      | 2.71               | 1.18      |
| Flow session 1                                                                                                                                                                                                                       | 3.22     | 0.48      | 3.29              | 0.50      | 3.16               | 0.46      |
| Flow session 2                                                                                                                                                                                                                       | 3.33     | 0.52      | 3.40              | 0.53      | 3.26               | 0.51      |
| Objective performance session 1                                                                                                                                                                                                      | 533      | 723       | 719               | 943       | 346                | 297       |
| Objective performance session 2                                                                                                                                                                                                      | 662      | 902       | 909               | 1,130     | 413                | 480       |
| Subjective performance session 1                                                                                                                                                                                                     | 2.63     | 1.16      | 2.76              | 1.22      | 2.50               | 1.08      |
| Subjective performance session 2                                                                                                                                                                                                     | 3.09     | 1.30      | 3.26              | 1.35      | 2.91               | 1.24      |
| Note. $N = 257$ ( $n_{\text{balance}} = 129$ ; $n_{\text{overload}} = 128$ ). Flow was measured on a 5-point and subjective performance on a 7-point response scale. For each variable, higher values show it to be more pronounced. |          |           |                   |           |                    |           |

Table 2  
*Spearman Correlations for Flow and Performance*

| Variable                                                                                                                                | 1      | 2      | 3      | 4      | 5      | 6 |
|-----------------------------------------------------------------------------------------------------------------------------------------|--------|--------|--------|--------|--------|---|
| 1. Flow session 1                                                                                                                       | –      |        |        |        |        |   |
| 2. Flow session 2                                                                                                                       | .58*** | –      |        |        |        |   |
| 3. Objective performance session 1                                                                                                      | .25*** | .21*** | –      |        |        |   |
| 4. Objective performance session 2                                                                                                      | .15*   | .22*** | .81*** | –      |        |   |
| 5. Subjective performance session 1                                                                                                     | .59*** | .41*** | .23*** | .17**  | –      |   |
| 6. Subjective performance session 2                                                                                                     | .39*** | .62*** | .21*** | .29*** | .69*** | – |
| Note. $N = 257$ ( $n_{\text{balance}} = 129$ ; $n_{\text{overload}} = 128$ ). The Holm method was employed to correct for multiplicity. |        |        |        |        |        |   |
| * $p < .05$ , ** $p < .01$ , *** $p < .001$                                                                                             |        |        |        |        |        |   |

Results regarding the mood induction are not the focus of this article and can be found in Supplement S2 and S3; in the following, only the results regarding the flow-performance relationship in the pooled sample are presented.

## Manipulation Check

Manipulation check results can be found in Supplement S2. To check whether the task difficulty manipulation yielded the intended differences in participants' experiences while playing Tetris, perceived skill-challenge balance was predicted by condition, session, and a random intercept per participant. The model ( $R^2_c = .64$ ,  $R^2_m = .05$ ) revealed main effects of condition,  $F(1, 256) = 7.31$ ,  $p = .007$ ,  $\eta_p^2 = .03$ , and of session,  $F(1, 256) = 32.78$ ,  $p < .001$ ,  $\eta_p^2 = .11$ , on perceived skill-challenge balance. As intended, skill-challenge balance was perceived to be more optimal in the balance condition as compared to the overload condition ( $\beta = .30$ , 95% CI [.08, .51]). Most likely due to practice effects, this was also the case for the second session as compared to the first ( $\beta = .31$ , 95% CI [.20, .41]). Thus, despite the small effect size of the condition effect, the task difficulty manipulation was overall successful.

## Multilevel Models

Multilevel model results are displayed in Table 3. The model testing H1a ( $R^2_c = .65$ ,  $R^2_m = .11$ ) revealed main effects of session,  $F(1, 263.04) = 4.73$ ,  $p = .031$ ,  $\eta_p^2 = .02$ , condition,  $F(1, 257.60) = 20.12$ ,  $p < .001$ ,  $\eta_p^2 = .07$ , and flow,  $F(1, 509.56) = 21.61$ ,  $p < .001$ ,  $\eta_p^2 = .04$ , on objective performance. Participants obtained better performance scores in the second session as compared to the first session ( $\beta = .12$ , 95% CI [.01, .22]) and in the balance condition as compared to the overload condition ( $\beta = .48$ , 95% CI [.27, .69]). Beyond these effects, flow was positively associated with performance scores ( $\beta = .20$ , 95% CI [.11, .28]), supporting H1a.

Table 3  
Multilevel Model Results

| Effect                                                                                                                                                                                                                                                                           | Objective performance models |       |        |        |        | Subjective performance models |          |      |        |      |        |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|-------|--------|--------|--------|-------------------------------|----------|------|--------|------|--------|------|
|                                                                                                                                                                                                                                                                                  | Estimate                     | SE    | 95% CI |        | p      | SD                            | Estimate | SE   | 95% CI |      | p      | SD   |
|                                                                                                                                                                                                                                                                                  |                              |       | LL     | UL     |        |                               |          |      | LL     | UL   |        |      |
| Model 1                                                                                                                                                                                                                                                                          |                              |       |        |        |        |                               |          |      |        |      |        |      |
| Fixed effects                                                                                                                                                                                                                                                                    |                              |       |        |        |        |                               |          |      |        |      |        |      |
| Intercept                                                                                                                                                                                                                                                                        | 354.04                       | 65.54 | 225.81 | 482.30 | < .001 |                               | 2.65     | 0.08 | 2.49   | 2.81 | < .001 |      |
| Session (session 2 vs. session 1)                                                                                                                                                                                                                                                | 95.02                        | 43.69 | 9.35   | 180.73 | .031   |                               | 0.31     | 0.05 | 0.21   | 0.41 | < .001 |      |
| Condition (balance vs. overload)                                                                                                                                                                                                                                                 | 390.63                       | 87.08 | 220.14 | 561.09 | < .001 |                               | 0.11     | 0.11 | – 0.11 | 0.33 | .334   |      |
| Flow                                                                                                                                                                                                                                                                             | 321.27                       | 69.10 | 186.04 | 456.41 | < .001 |                               | 1.41     | 0.08 | 1.25   | 1.58 | < .001 |      |
| Random effects                                                                                                                                                                                                                                                                   |                              |       |        |        |        |                               |          |      |        |      |        |      |
| Intercept                                                                                                                                                                                                                                                                        |                              |       |        |        |        | 601.90                        |          |      |        |      |        | 0.80 |
| Residual                                                                                                                                                                                                                                                                         |                              |       |        |        |        | 488.40                        |          |      |        |      |        | 0.56 |
| Model 2                                                                                                                                                                                                                                                                          |                              |       |        |        |        |                               |          |      |        |      |        |      |
| Fixed effects                                                                                                                                                                                                                                                                    |                              |       |        |        |        |                               |          |      |        |      |        |      |
| Intercept                                                                                                                                                                                                                                                                        | 353.75                       | 65.60 | 225.51 | 481.99 | < .001 |                               | 2.65     | 0.08 | 2.49   | 2.82 | < .001 |      |
| Session (session 2 vs. session 1)                                                                                                                                                                                                                                                | 92.55                        | 44.36 | 5.63   | 179.48 | .038   |                               | 0.32     | 0.05 | 0.22   | 0.42 | < .001 |      |
| Condition (balance vs. overload)                                                                                                                                                                                                                                                 | 393.67                       | 87.64 | 222.26 | 565.07 | < .001 |                               | 0.09     | 0.11 | – 0.13 | 0.31 | .414   |      |
| Flow (within-person)                                                                                                                                                                                                                                                             | 344.76                       | 99.19 | 150.39 | 539.14 | < .001 |                               | 1.31     | 0.11 | 1.09   | 1.53 | < .001 |      |
| Flow (between-person)                                                                                                                                                                                                                                                            | 299.03                       | 96.49 | 110.32 | 487.74 | .002   |                               | 1.53     | 0.12 | 1.29   | 1.77 | < .001 |      |
| Random effects                                                                                                                                                                                                                                                                   |                              |       |        |        |        |                               |          |      |        |      |        |      |
| Intercept                                                                                                                                                                                                                                                                        |                              |       |        |        |        | 602.40                        |          |      |        |      |        | 0.80 |
| Residual                                                                                                                                                                                                                                                                         |                              |       |        |        |        | 488.80                        |          |      |        |      |        | 0.56 |
| Note. N = 257. Estimates are unstandardized coefficients. Flow is grand-mean centered. Flow (between-person) is participants' grand-mean centered mean flow and Flow (within person) is person-mean centered flow. CI = confidence interval; LL = lower limit; UL = upper limit. |                              |       |        |        |        |                               |          |      |        |      |        |      |

The model testing H1b ( $R^2_c = .79$ ,  $R^2_m = .38$ ) showed a similar pattern of results for subjective performance, except that only the main effects of session,  $F(1, 262.13) = 37.38$ ,  $p < .001$ ,  $\eta_p^2 = .12$ , and flow,  $F(1, 506.62) = 283.43$ ,  $p < .001$ ,  $\eta_p^2 = .36$ , were significant, but not the main effect of condition,  $F(1, 256.86) = 0.94$ ,  $p = .334$ ,  $\eta_p^2 = .00$ . Participants evaluated their performance more positively in the second session as compared to the first session ( $\beta = .25$ , 95% CI [.17, .32]) but similar across conditions ( $\beta = .09$ , 95% CI [−.09, .26]). More positive performance evaluations were further associated with more intense flow ( $\beta = .57$ , 95% CI [.50, .64]). Therefore, H1b was supported as well.

Decomposing the flow predictor in the model testing H2a ( $R^2_c = .64$ ,  $R^2_m = .11$ ) yielded significant main effects of session  $F(1, 255) = 4.35$ ,  $p = .038$ ,  $\eta_p^2 = .02$ , condition,  $F(1, 254) = 20.18$ ,  $p < .001$ ,  $\eta_p^2 = .07$ , and of both the within-person flow component,  $F(1, 255) = 12.08$ ,  $p < .001$ ,  $\eta_p^2 = .05$ , and the between-person flow component,  $F(1, 254) = 9.61$ ,  $p = .002$ ,  $\eta_p^2 = .04$ , on objective performance (see Fig. 1). The results reaffirmed that participants obtained higher performance scores in the second session as compared to the first session ( $\beta = .11$ , 95% CI [.01, .22]) and in the balance condition as compared to the overload condition ( $\beta = .48$ , 95% CI [.27, .69]). Additionally, they showed that both within the same participant ( $\beta = .09$ , 95% CI [.04, .15]) and when drawing comparisons between participants ( $\beta = .17$ , 95% CI [.06, .27]), more intense flow was linked to higher performance scores, which supports H2a.

The corresponding model testing H2b ( $R^2_c = .80$ ,  $R^2_m = .40$ ) for subjective performance showed similar results (see Fig. 1), as significant main effects emerged for session,  $F(1, 255) = 39.01$ ,  $p < .001$ ,  $\eta_p^2 = .13$ , and for both the within-person flow component,  $F(1, 255) = 132.70$ ,  $p < .001$ ,  $\eta_p^2 = .34$ , and between-person flow component,  $F(1, 254) = 152.79$ ,  $p < .001$ ,  $\eta_p^2 = .38$ , but not for condition,  $F(1, 254) = 0.67$ ,  $p = .414$ ,  $\eta_p^2 = .00$ . Again, the result reaffirmed that participants evaluated their performance more positively in the second session than in the first ( $\beta = .25$ , 95% CI [.17, .33]), but similar across the balance and overload conditions ( $\beta = .07$ , 95% CI [−.10, .25]). Moreover, more intense flow came with more positive evaluations of one's performance not only within the same participant ( $\beta = .23$ , 95% CI [.19, .27]), but also when comparing between different participants ( $\beta = .55$ , 95% CI [.47, .64]), which further supports H2b.

Sensitivity analyses (see Supplement S4), excluding influential data points and outliers as well as controlling for Tetris-related background variables and mood induction group, confirmed the multilevel model results' robustness. Only for objective performance, the between-person flow component was no longer significant when 30 level 2 outliers (all very high scores) were removed and when participants' Tetris-related background was controlled for.

## Multilevel Mediation

Figure 2 shows the multilevel mediation model results for each objective and subjective performance. Flow partially mediated the effect of task difficulty condition on objective performance, with the balance condition increasing flow and thereby indirectly also performance scores by 43.78 points, 95% CI [7.65, 87.81]. Moreover, for subjective performance, a significant indirect effect of task difficulty through flow emerged, although the total effect of task difficulty on subjective performance was not significant. The more intense flow in the balance condition indirectly increased participants' performance ratings by 0.19 points, 95% CI [0.04, 0.35]. Thus, H3a was supported, and H3b was partially supported.

Sensitivity analyses (see Supplement S5), excluding influential data points and outliers as well as controlling for Tetris-related background variables and mood induction group, showed these results to be robust, except for the indirect effect of task difficulty on objective performance through flow becoming non-significant when 30 outliers (all very high scores) were excluded.

## Lagged Regression Models

Lagged regression model results are presented in Table 4. The model testing H4a,  $F(3, 253) = 9.94$ ,  $p < .001$ , explained 43% of the variance in objective performance in the second session ( $R^2 = .43$ ). Objective performance in the first session was a strong positive predictor, suggesting relative temporal stability. Condition was a significant predictor as well, with objective performance in the second session being better in the balance than in the overload condition. When controlling for these factors, flow in the first session did not predict performance in the second session, so that H4a was not supported.

Table 4  
*Lagged Regression Model Results*

| Effect                                                                                                                                                                                                                                             | <b>Objective performance models</b> |        |         |        |          | <b>Subjective performance models</b> |      |        |      |          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|--------|---------|--------|----------|--------------------------------------|------|--------|------|----------|
|                                                                                                                                                                                                                                                    | Estimate                            | SE     | 95% CI  |        | <i>p</i> | Estimate                             | SE   | 95% CI |      | <i>p</i> |
|                                                                                                                                                                                                                                                    |                                     |        | LL      | UL     |          |                                      |      | LL     | UL   |          |
| Intercept                                                                                                                                                                                                                                          | −44.73                              | 366.64 | −766.79 | 677.32 | .903     | 0.96                                 | 0.39 | 0.20   | 1.73 | .014     |
| Condition (balance vs. overload)                                                                                                                                                                                                                   | 207.27                              | 85.47  | 38.95   | 375.59 | .016     | 0.14                                 | 0.12 | −0.10  | 0.29 | .247     |
| Flow session 1                                                                                                                                                                                                                                     | 61.91                               | 115.98 | −166.50 | 290.32 | .594     | 0.01                                 | 0.14 | −0.27  | 0.38 | .940     |
| Performance session 1                                                                                                                                                                                                                              | 0.76                                | 0.29   | 0.19    | 1.32   | .009     | 0.77                                 | 0.07 | 0.64   | 0.90 | < .001   |
| <i>Note.</i> <i>N</i> = 257. Estimates are unstandardized coefficients. Heteroscedasticity-consistent (HC3) standard errors and corresponding 95% confidence intervals are reported. CI = confidence interval; LL = lower limit; UL = upper limit. |                                     |        |         |        |          |                                      |      |        |      |          |

Similar results were obtained for the model testing H4b,  $F(3, 253) = 78.15$ ,  $p < .001$ , which explained 48% of the variance in subjective performance in the second session ( $R^2 = .48$ ). While subjective performance in the first session was a strong predictor, neither condition nor flow in the first session were significant. Hence, H4b was not supported either.

Sensitivity analyses (see Supplement S6), excluding influential data points and outliers, removing condition as a predictor, and controlling for Tetris-related background variables as well as for mood induction group, did not yield substantially different results for each model. However, excluding 24 objective performance outliers in the second session (all very high scores) led to flow in the first session becoming a significant predictor. Moreover, the exploratory analysis found no evidence for a reverse relationship between performance and flow, as performance in the first session did not predict flow in the second session beyond baseline and condition (see Supplement S7).

## Discussion

The current study was aimed at disentangling the flow-performance relationship, by examining flow and subjective as well as objective performance in an established flow-eliciting task, the computer game Tetris, across two different task difficulty levels and two sessions. The findings not only corroborate that flow experiences are positively associated with performance (Harris et al., 2023; Jinmin & Qi, 2023), but also give deeper insights into the nature of this relationship: Together, the findings suggest that being in a state of flow and performing well are closely coupled phenomena, especially, but not exclusively, in terms of having a positive subjective task experience. Performance enhancements in association with flow seem to occur both relative to oneself and to other individuals, but no evidence was found that this extends into subsequent task execution. Task difficulty had causal effects on both flow and performance, with flow being one potential pathway through which task difficulty positively affects performance. Moreover, flow and performance showed a robust positive association that remained stable beyond contextual task conditions including difficulty, practice, or interindividual differences. Thus, experiencing flow appears to play a distinct role in the effective task engagement that accompanies successful task performance.

Our first set of multilevel models tested the association between flow and performance while accounting for other factors that are relevant for task execution, such as the task's difficulty level, participants' learning curve, and stable differences across participants. More intense flow predicted both better objective performance scores and more favorable subjective performance ratings beyond task difficulty, practice effects, and interindividual differences. This result shows that the positive association between flow and performance is not merely a by-product of these other task conditions facilitating both flow and performance at the same time. While the most favorable conditions for participants to perform well were under balanced task difficulty and when having had some time to practice in the second session, participants still differed in their performance, and part of this variance was explained by differences in flow: When comparing participants who faced these same task conditions, the ones reporting more flow also tended to be the ones who performed better. This is an important insight insofar as flow-conducive task conditions, specifically skill-challenge balance, are also generally seen as favorable for performance (Barthelmäs & Keller, 2021), and previous evidence suggested that this alone might account for an

apparent positive association between flow and performance (Keller & Bless, 2008; Keller & Blomann, 2008). However, since flow explained additional variance in performance, favorable task conditions do not seem to be the sole driving factor behind the flow-performance relationship. Instead, the flow experience plays a distinct role in successful task performance.

The second set of multilevel models examined the association between flow and performance in more detail by separating fluctuations within the same participant across sessions from differences between participants. Decomposing flow into a within-person and between-person component revealed that the positive association between flow and performance could be found on the individual as well as on the interindividual level, suggesting that: a) experiencing more flow than usual comes with an increased likelihood to perform better, and b) experiencing more flow than others comes with an increased likelihood to also outperform them. Despite similar effect sizes, sensitivity analyses showed the flow-performance relationship to be more robust on the individual level. One could speculate that this is why some studies that focused primarily on the interindividual level did not find a significant flow-performance relationship (Schiefele & Roussakis, 2006; Schüler & Brunner, 2009). Importantly, this result highlights that the flow-performance relationship is not only driven by stable differences across individuals, and thus also not only by highly skilled performers. Further examining fluctuations in flow and performance within individuals seems to be a promising approach to get closer insights into the flow-performance relationship. Previous research has found that flow is a highly dynamic experience (Ceja & Navarro, 2011; Fullagar & Kelloway, 2009). The daily variability in flow has been shown to moderate overall flow-performance associations (Stollberger & Debus, 2020) and to predict daily variability in performance as well as to mediate effects of other performance-relevant variables such as intrusions (Aitken et al., 2023) or multitasking (Peifer & Zipp, 2019). However, to the best of our knowledge, these studies have not decomposed flow into within-person and between-person components. Future studies should therefore look further into decomposing the flow-performance relationship, to gain a better understanding of its stable versus dynamic aspects.

As Harris et al. (2023) point out, several mechanisms through which flow and performance could be connected have been discussed in the literature, but for the most part have not been formally tested, yet. One plausible explanation for the link between flow and performance is provided by our mediation results, which suggest that task difficulty affects performance, at least in part, through flow. For objective performance, this mediation was partial, meaning that facilitating flow was only one possible way in which the balanced task difficulty condition could allow participants to perform better. This result is not surprising, insofar as balanced task difficulty is in itself a favorable condition for good performance (Barthelmäs & Keller, 2021)—but importantly, experiencing flow adds to this effect. For subjective performance, on the other hand, task difficulty had no significant total effect, but a significant indirect effect through flow. Thus, participants seem to have based their performance evaluations primarily on emergent experiences during the task, rather than directly considering objective task conditions. It seems plausible that experiencing flow serves as a heuristic based on which individuals infer that they are performing well. Taken together, the mediation results suggest that intense flow and excellent performance may frequently co-occur because they can both arise from balanced task difficulty, and that the flow experience is a potential pathway through which part of the performance enhancements that result from balanced task difficulty manifest. However, the mediation results need to be interpreted with caution, as the retrospective measurement of flow renders the distinction between mediator and outcome purely theoretical, meaning that performance could also be the mediator for task difficulty effects on flow.

Despite our results consistently showing a close positive link between flow and performance, we found no evidence that this relationship extends beyond immediate task execution. As flow in the first session did not predict performance in the second session when controlling for baseline and condition, nor vice versa, intense flow and enhanced performance seem to be closely coupled during immediate task execution but do not seem to drive one another beyond this concurrent relationship. It seems plausible that flow and performance are closely coupled in the immediate execution of a task because they reflect shared task engagement processes: For instance, more effective attentional processing is linked not only to experiencing flow (de Sampaio Barros et al., 2018; Harris et al., 2017; Marty-Dugas et al., 2021; Thissen & Oettingen, 2026) but should at the same time benefit the execution of a task like Tetris, a game that predominantly relies on visuospatial working memory (Lau-Zhu et al., 2017). However, which task engagement processes, specifically, manifest in both flow experiences and performance enhancements may be task-dependent and needs to be investigated by future research. Further studies could also examine how flow and performance co-develop in the long term, as the absence of positive short-term carry-over effects does not rule out indirect performance benefits through flow over a longer period of time. In fact, Schüler and Brunner (2009) found flow to positively predict training motivation in marathon runners, which should, in turn, benefit their future race performances. Thus, beyond the concurrent coupling of flow and performance that we observed in this study, the key benefit of flow for performance might, in fact, lie in its long-term implication for intrinsic motivation.

One should note that, in line with the notion of flow as an intrinsically motivating experience, throughout our analyses, associations between flow and subjective performance were stronger and more robust than between flow and objective performance. As both

performance indicators were only moderately correlated, performing a somewhat challenging task well does not always seem to translate into experiencing a sense of competency, and vice versa. However, feeling competent during the execution of such a task appears to often be accompanied by a flow experience. This combination of perceived competency with full task engagement could be the psychological process by which flow states become a catalyst for intrinsic motivation, which should, in turn, also benefit objective performance outcomes to some degree.

## Limitations

The current study has several limitations in that it is a secondary data analysis, which relies on only two measurement points for multilevel modeling and retrospective measurements of both flow and subjective performance. Therefore, more complex within-person processes could not be modeled, and the temporal ordering of the mediation model remains somewhat speculative. Moreover, even though using Tetris as the experimental task in a laboratory environment allowed for highly controlled conditions, the findings might not apply to more sophisticated real-world performance situations. Examining other tasks than Tetris and other populations than our student sample could produce divergent results, so future studies should test the generalizability of our findings.

## Conclusion

Our findings suggest that flow experiences accompany better performance during task execution, both on a subjective and, to a lesser degree, also on an objective level. This association remained significant when accounting for other factors that appear relevant for both flow and performance, including task difficulty, practice, and individual differences. Moreover, this positive relationship was not only driven by differences between individuals, which could be confounded by stable third variables such as individual skill levels, but also by within-person fluctuations in flow. Balancing task difficulty with participants' skill levels facilitates both flow and performance, and it seems plausible that part of this performance enhancement occurs indirectly through flow. However, no evidence for carry-over effects between performance and flow was found, so that, at least in the short term, the flow-performance relationship may be best described as a concurrent coupling.

## Declarations

The authors have no relevant financial or non-financial interests to disclose.

## Author Contribution

B.T. designed the study, collected the data, performed the statistical analyses, wrote the main manuscript text, and prepared tables and figures. M.L. helped design the study, coded the game used as the experimental task, and reviewed the manuscript. B.K. helped collect the data, transcribed and cleaned the data in preparation for the statistical analyses, and reviewed the manuscript. M.B. helped design the study and reviewed the manuscript. G.O. helped design the study and reviewed the manuscript.

## Acknowledgement

The authors would like to thank Franziska Rothering, Lilli Fohrer, Danny Lightfoot, and Lesley Tan for their help with data transcriptions.

## Data Availability

The data and code that supports the current findings were deposited into the Open Science Framework database and are available at the following URL: [https://osf.io/r9ahs/overview?view\\_only=fd400acae7394c9183e5e0bd39730a2d](https://osf.io/r9ahs/overview?view_only=fd400acae7394c9183e5e0bd39730a2d).

## References

1. Aitken, J. A., Kaplan, S. A., & Kuykendall, L. (2023). Going with(out) the flow at home: Day-to-day variability in flow and performance while teleworking. *European Journal of Work and Organizational Psychology*, 32(5), 662–677.  
<https://doi.org/10.1080/1359432X.2023.2211271>

2. Bakker, A. B., Oerlemans, W., Demerouti, E., Slot, B. B., & Ali, D. K. (2011). Flow and performance: A study among talented Dutch soccer players. *Psychology of Sport and Exercise*, 12(4), 442–450. <https://doi.org/10.1016/j.psychsport.2011.02.003>
3. Barthelmäs, M., & Keller, J. (2021). Antecedents, boundary conditions and consequences of flow. In C. Peifer & S. Engeser (Eds.), *Advances in flow research* (pp. 71–107). Springer International Publishing. [https://doi.org/10.1007/978-3-030-53468-4\\_3](https://doi.org/10.1007/978-3-030-53468-4_3)
4. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
5. Ceja, L., & Navarro, J. (2011). Dynamic patterns of flow in the workplace: Characterizing within-individual variability using a complexity science approach. *Journal of Organizational Behavior*, 32(4), 627–651. <https://doi.org/10.1002/job.747>
6. Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, 140(4), 980–1008. <https://doi.org/10.1037/a0035661>
7. Christandl, F., Mierke, K., & Peifer, C. (2018). Time flows: Manipulations of subjective time progression affect recalled flow and performance in a subsequent task. *Journal of Experimental Social Psychology*, 74, 246–256. <https://doi.org/10.1016/j.jesp.2017.09.015>
8. Cseh, G. M., Phillips, L. H., & Pearson, D. G. (2015). Flow, affect and visual creativity. *Cognition and Emotion*, 29(2), 281–291. <https://doi.org/10.1080/02699931.2014.913553>
9. Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety* (1 ed.). Jossey-Bass.
10. de Sampaio Barros, M. F., Araújo-Moreira, F. M., Trevelin, L. C., & Radel, R. (2018). Flow experience and the mobilization of attentional resources. *Cognitive Affective & Behavioral Neuroscience*, 18(4), 810–823. <https://doi.org/10.3758/s13415-018-0606-4>
11. Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, 32(3), 158–172. <https://doi.org/10.1007/s11031-008-9102-4>
12. 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>
13. Fullagar, C. J., & Kelloway, E. K. (2009). Flow at work: An experience sampling approach. *Journal of Occupational and Organizational Psychology*, 82(3), 595–615. <https://doi.org/10.1348/096317908X357903>
14. Fullagar, C. J., Knight, P. A., & Sovern, H. S. (2013). Challenge/skill balance, flow, and performance anxiety. *Applied Psychology*, 62(2), 236–259. <https://doi.org/10.1111/j.1464-0597.2012.00494.x>
15. Gaggioli, A., Chirico, A., Mazzoni, E., Milani, L., & Riva, G. (2017). Networked Flow in musical bands. *Psychology of Music*, 45(2), 283–297. <https://doi.org/10.1177/0305735616665003>
16. Green, P., & MacLeod, C. J. (2016). Simr: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
17. Harris, D. J., Allen, K. L., Vine, S. J., & Wilson, M. R. (2023). A systematic review and meta-analysis of the relationship between flow states and performance. *International Review of Sport and Exercise Psychology*, 14(1), 1–29. <https://doi.org/10.1080/1750984X.2021.1929402>
18. Harris, D. J., Vine, S. J., & Wilson, M. R. (2017). Is flow really effortless? The complex role of effortful attention. *Sport Exercise and Performance Psychology*, 6(1), 103–114. <https://doi.org/10.1037/spy0000083>
19. 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 and Exercise Psychology*, 24(2), 133–150. <https://doi.org/10.1123/jsep.24.2.133>
20. Jinmin, Z., & Qi, F. (2023). Relationship between learning flow and academic performance among students: A systematic evaluation and meta-analysis. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1270642>. Article e1270642.
21. Keller, J., & Bless, H. (2008). Flow and regulatory compatibility: An experimental approach to the flow model of intrinsic motivation. *Personality and Social Psychology Bulletin*, 34(2), 196–209. <https://doi.org/10.1177/0146167207310026>
22. Keller, J., & Blomann, F. (2008). Locus of control and the flow experience: An experimental analysis. *European Journal of Personality*, 22(7), 589–607. <https://doi.org/10.1002/per.692>
23. Keller, J., Ringelhan, S., & Blomann, F. (2011). Does skills–demands compatibility result in intrinsic motivation? Experimental test of a basic notion proposed in the theory of flow-experiences. *The Journal of Positive Psychology*, 6(5), 408–417. <https://doi.org/10.1080/17439760.2011.604041>
24. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package.
25. Tests in linear mixed effects models *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>

26. Lau-Zhu, A., Holmes, E. A., Butterfield, S., & Holmes, J. (2017). Selective association between Tetris game play and visuospatial working memory: A preliminary investigation. *Applied Cognitive Psychology*, 31(4), 438–445. <https://doi.org/10.1002/acp.3339>
27. Lavoie, R. V., & Main, K. J. (2024). What happens when flow ends? How and why your creativity is limited after a flow experience. *Current Psychology*, 43(41), 32128–32142. <https://doi.org/10.1007/s12144-024-06591-4>
28. Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, 7(5), 573–579. <https://doi.org/10.1111/2041-210X.12512>
29. Liu, W., van der Linden, D., & Bakker, A. B. (2022). Strengths use and work-related flow: An experience sampling study on implications for risk taking and attentional behaviors. *Journal of Managerial Psychology*, 37(1), 47–60. <https://doi.org/10.1108/JMP-07-2020-0403>
30. Lu, H., Linden, D. V. D., & Bakker, A. B. (2025). Disentangling the effects of task difficulty and effort on flow experience. *Psychological Research Psychologische Forschung*, 89(4). <https://doi.org/10.1007/s00426-025-02128-x>. Article e113.
31. Marty-Dugas, J., Howes, L., & Smilek, D. (2021). Sustained attention and the experience of flow. *Psychological Research Psychologische Forschung*, 85, 2682–2696. <https://doi.org/10.1007/s00426-020-01433-x>
32. Mehta, P., & Vyas, M. (2022). A systematic literature review on the experience of flow and its relation to intrinsic motivation in students. *Indian Journal of Positive Psychology*, 13(3), 299–304. <https://doi.org/10.15614/ijpp%2F2022%2Fv13i3%2F215403>
33. Norsworthy, C., Dimmock, J. A., Miller, D. J., Krause, A., & Jackson, B. (2023). 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. *International Journal of Applied Positive Psychology*, 8(2), 309–337. <https://doi.org/10.1007/s41042-023-00092-8>
34. Palomäki, J., Tammi, T., Lehtonen, N., Seittenranta, N., Laakasuo, M., Abuhamdeh, S., Lappi, O., & Cowley, B. U. (2021). The link between flow and performance is moderated by task experience. *Computers in Human Behavior*, 124, e106891. <https://doi.org/10.1016/j.chb.2021.106891>
35. Pajitnov, A. L. (1984). *Tetris [Computer game]*. Tetris Company.
36. Peifer, C., & Zipp, G. (2019). All at once? The effects of multitasking behavior on flow and subjective performance. *European Journal of Work and Organizational Psychology*, 28(5), 682–690. <https://doi.org/10.1080/1359432X.2019.1647168>
37. R Core Team. (2025). *R: A Language and Environment for Statistical Computing (Version 4.5.1) [Computer software]*. R Foundation for Statistical Computing. <https://www.R-project.org>
38. Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow-Erlebens [The assesement of flow experiences]. In J. Stiensmeyer-Pelster & F. Rheinberg (Eds.), *Diagnostik von Motivation und Selbst-Konzept [Diagnostics of motivation and self-concept]* (pp. 261–279). Hogrefe. <https://doi.org/10.23668/psycharchives.8590>
39. Salanova, M., Rodríguez-Sánchez, A. M., Schaufeli, W. B., & Cifre, E. (2014). Flowing together: A longitudinal study of collective efficacy and collective flow among workgroups. *The Journal of Psychology*, 148(4), 435–455. <https://doi.org/10.1080/00223980.2013.806290>
40. Schiefele, U., & Roussakis, E. (2006). Die Bedingungen des Flow-Erlebens in einer experimentellen Spielsituation [The conditions of flow in an experimental gaming task]. *Journal of Psychology*, 214(4), 207–219. <https://doi.org/10.1026/0044-3409.214.4.207>
41. Schüler, J., & Brunner, S. (2009). The rewarding effect of flow experience on performance in a marathon race. *Psychology of Sport and Exercise*, 10(1), 168–174. <https://doi.org/10.1016/j.psychsport.2008.07.001>
42. Stoll, O., & Lau, A. (2005). Flow-Erleben beim Marathonlauf [Flow experience during marathon running]. *Zeitschrift für Sportpsychologie*, 12(3), 75–82. <https://doi.org/10.1026/1612-5010.12.3.75>
43. Stollberger, J., & Debus, M. E. (2020). Go with the flow, but keep it stable? The role of flow variability in the context of daily flow experiences and daily creative performance. *Work & Stress*, 34(4), 342–358. <https://doi.org/10.1080/02678373.2019.1695293>
44. Thissen, B. A. K., & Oettingen, G. (2026). How optimal is the optimal experience? Toward a more nuanced understanding of the relationship between flow states, attentional performance, and perceived effort. *Psychology of Consciousness: Theory Research and Practice*, 13(2), 219–231. <https://doi.org/10.1037/cns0000398>
45. Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>

## Figures

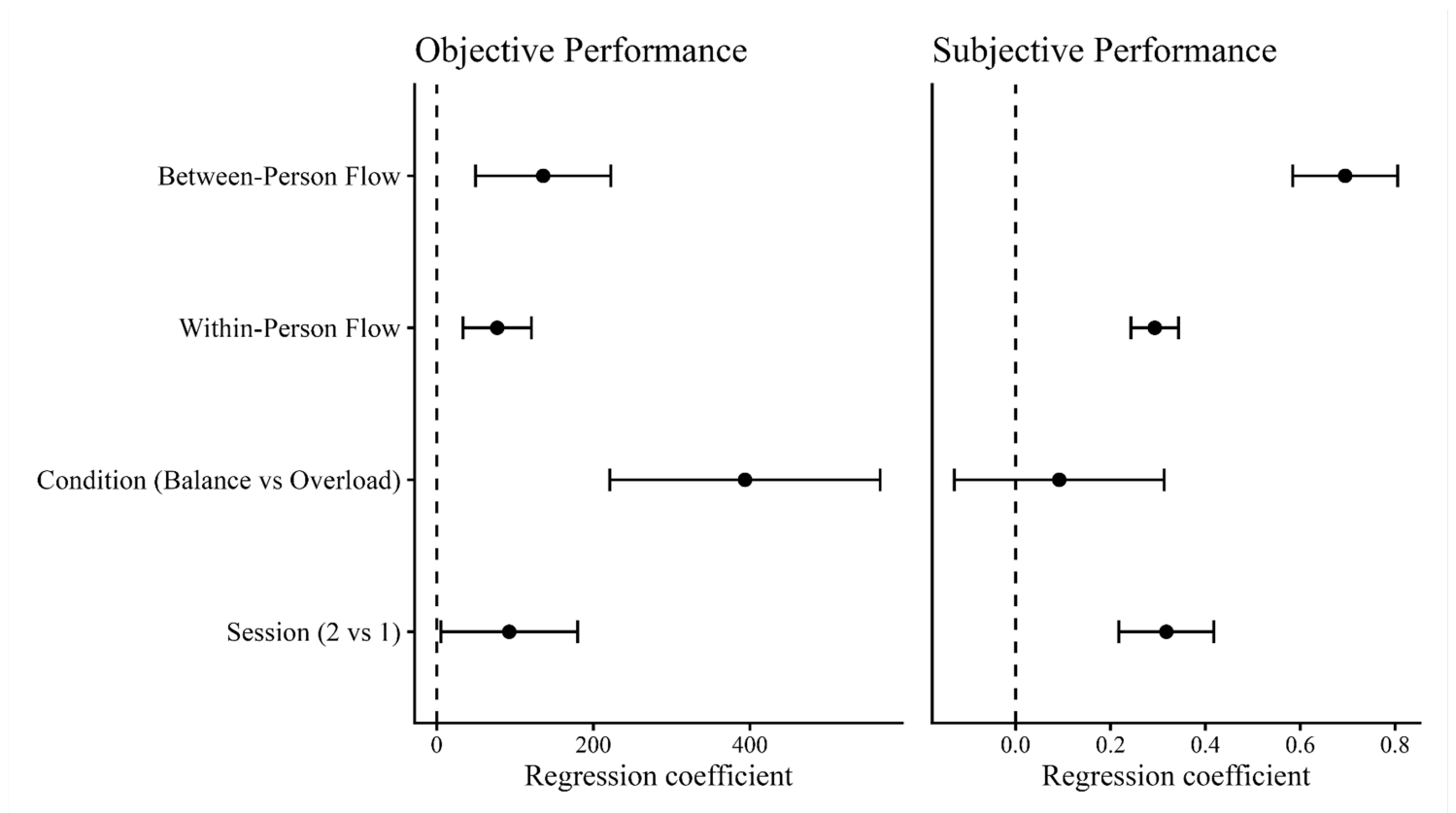
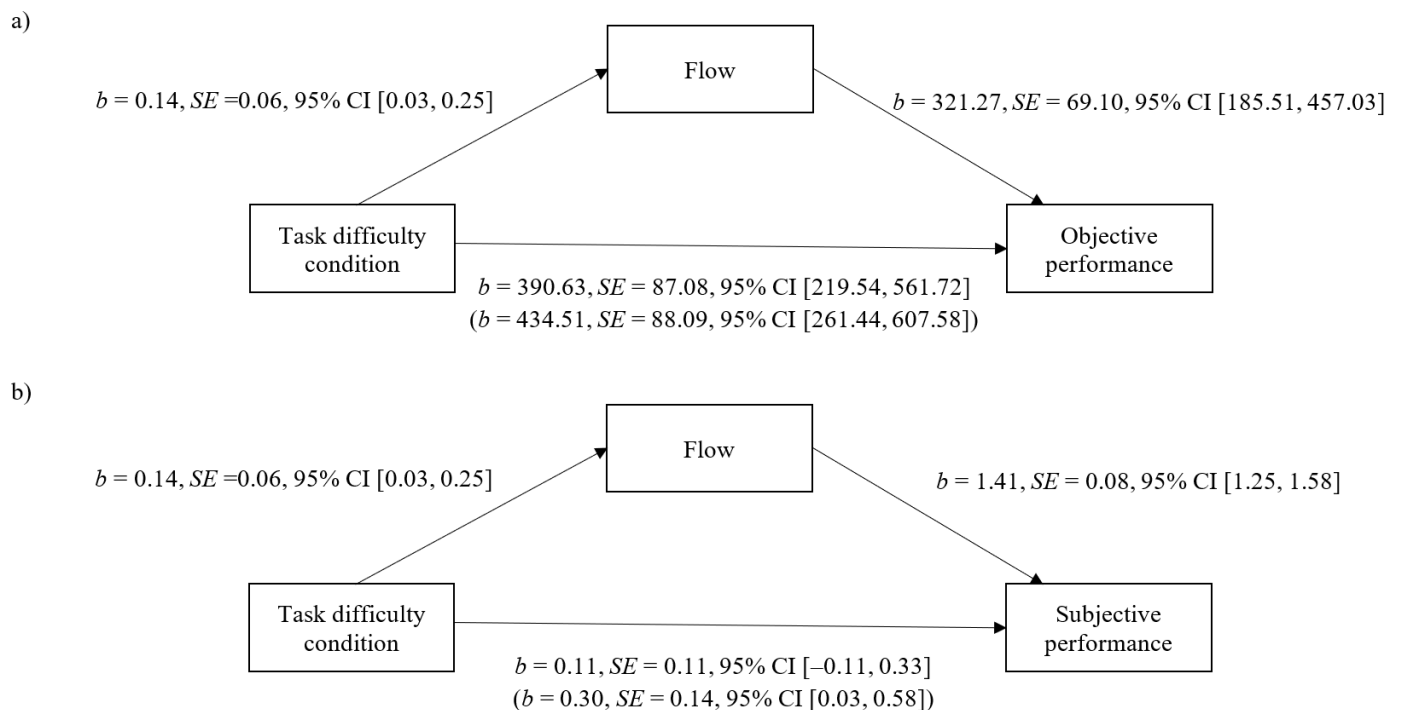


**Figure 1**

*Multilevel Model Results when Decomposing Flow into Within-Person and Between Person Components*

*Note.*  $N = 257$ . Points represent unstandardized regression coefficients. Between-person flow is participants' grand-mean centered mean flow and within person flow is person-mean centered flow. Error bars represent 95% confidence intervals. The dashed vertical line indicates a regression coefficient of zero.



## Figure 2

### *Multilevel Mediation Model Results*

Note.  $N = 257$ . Multilevel mediation models examining the indirect effect of condition (balance vs. overload) on performance through flow. Estimates are unstandardized coefficients. The coefficient in parentheses denotes the total effect of condition on performance before inclusion of the mediator. Significance is indicated by the CI excluding zero.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementHowFarDoesFlowGetYou.docx](#)