Balancing the Impossible: Understanding Human Performance in Multitasking Challenges

Sumin Kim
School of Computer Science and Engineering
Chung-Ang University
Republic of Korea
crapy0001@cau.ac.kr

Htet Arkar
School of Computer Science and Engineering
Chung-Ang University
Republic of Korea
hak3601@cau.ac.kr

Suho Moon
School of Computer Science and Engineering
Chung-Ang University
Republic of Korea
moonwalk725@cau.ac.kr

Yejin Jang
School of Computer Science and Engineering
Chung-Ang University
Republic of Korea
yejin923@cau.ac.kr

Ennji Park*
Interaction+X Lab
Chung-Ang University
Republic of Korea
enujipark@cau.ac.kr

Figure 1: Dual-Task Balancing Game

Abstract

Operationalizing the Index of Difficulty (ID) in real-time, multitasking contexts remains an ongoing challenge, largely due to the absence of standardized methods to model user behavior. This study addresses this gap by introducing a novel Fitts' law-based user model that integrates user-specific attributes and task-centric variables. To empirically validate this framework, a user study was conducted under six experimental conditions, combining two ball velocities with three plate radii. Task completion times were analyzed using repeated measures ANOVA, revealing that the proposed model not only successfully operationalizes ID but also pinpoints critical parameters to optimize task difficulty. The resulting predictive model demonstrates robust explanatory power, achieving an R2 score of 0.76.

CCS Concepts

• **Human-centered computing → User studies.**

Keywords

Multitasking Performance, Adaptive Interfaces, Task Difficulty

ACM Reference Format:
Sumin Kim, Htet Arkar, Suho Moon, Yejin Jang, and Ennji Park. 2024. Balancing the Impossible: Understanding Human Performance in Multitasking Challenges. In *Proceedings of Human Interface Media User Study (HIMUS '24)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Frequent multitasking scenarios are ubiquitous in contemporary life, requiring rapid task-switching and flexible cognitive strategies. Although individuals can enhance their multitasking skills through targeted gameplay, determining the most appropriate difficulty level for such training games remains a persistent challenge.

In response to this need, our study focuses on a real-time balancing game designed to cultivate quick reactions and effective task management. Specifically, players must manipulate two rotating discs to keep balls from falling off, thereby maintaining equilibrium for as long as possible.

Over the past decade, a variety of dual-task paradigms have been employed to investigate human performance under multi-component demands. While these studies have illuminated numerous factors—such as cognitive load, physical coordination, and user-specific attributes—that influence performance, significant gaps remain in modeling and predicting the Index of Difficulty (ID) within real-time, dynamic multitasking contexts. To address this lacuna, the present research integrates game parameters and human characteristics through a Fitts' law-based framework to formulate a predictive ID. By doing so, we aim to systematically guide difficulty adjustments in a real-time balancing task.

Unlike earlier work that utilized automated attention management systems, our experiment adopts a no-supervisor mode, thereby eliminating external guidance and permitting participants to self-regulate their task-switching strategies. Two main factors—ball velocity (high vs. low) and plate radius (small, medium, large)—were selected to yield six distinct task configurations. A within-subject design was implemented to ensure that each participant experienced all configurations, with counterbalancing to mitigate learning effects and sequence biases. Task completion time was employed as the principal performance metric, reflecting each participant's capacity to maintain equilibrium under varying conditions.

Building on this approach, the study is guided by the following central research questions:

- **How do variations in ball velocity and plate radius influence task completion time in a multitasking environment?**
- **What interactions, if any, exist between velocity and radius in determining task performance?**

Addressing these questions provides a robust examination of how systematically manipulating game parameters can optimize difficulty. Moreover, our predictive model—achieving an R2 score of 0.76—demonstrates strong potential in identifying parameter configurations that align with a targeted performance duration. Ultimately, these findings contribute to a deeper understanding of human multitasking dynamics and inform the development of adaptive systems that fine-tune difficulty levels based on individual user characteristics, thereby enhancing both skill acquisition and overall performance in complex, multitasking scenarios.

2 Related Work

Research on multitasking consistently underscores challenges associated with managing multiple tasks simultaneously, such as switch costs and interruptions to cognitive flow, both of which adversely affect performance. Theoretical frameworks, including Threaded Cognition, conceptualize human attention as a finite resource allocated among concurrent tasks, thereby providing explanations for the limitations of multitasking.

Attention Management Systems (AMSs) have emerged as a potential solution, employing strategies such as interruption timing and resumption cues to alleviate multitasking inefficiencies. While AMSs have shown promise, their heuristic-based and ACT-R models have encountered difficulties in generalizing across diverse user behaviors and application domains. More recent approaches utilizing reinforcement learning (RL) allow AMSs to adapt task-switching policies dynamically. However, challenges remain in translating these advances into robust, real-world scenarios.

Experimental paradigms like the Dual-Task Balancing Game offer a controlled environment to examine multitasking performance. Previous studies integrating RL-based AMSs have demonstrated improved outcomes and reduced workload. In contrast, the present study removes supervisory interventions to focus exclusively on human performance under varying task parameters, thereby expanding our understanding of the fundamental constraints and capabilities of human multitasking.

3 Methodology

This research adopted a multifaceted approach to define and validate an Index of Difficulty (ID) that integrates both game-specific parameters and human performance metrics. The ID was designed to reflect the interplay between task complexity and individual ability, thereby capturing a comprehensive view of user performance in a real-time balancing game.

3.1 Index of Difficulty (ID) Formulation

We operationalized the ID by combining game-related factors (plate radius and minimum ball velocity) with human performance measures (task time and reaction time). Specifically, the ID was defined as

$$\mathrm{ID} = \log_2\left(\frac{V}{R} + 1\right) \times \frac{\Delta t}{\mu_{\Delta t}} \times \frac{T}{\mu_T} \tag{1}$$

where R denotes the plate radius, V is the minimum velocity of the balls, Δt is the time spent on the task before switching, T is the reaction time to visual stimuli, $\mu_{\Delta t}$ and μ_T represent the mean values of Δt and T across all participants, respectively.

This formulation includes a logarithmic term capturing the ratio of velocity to radius, thereby integrating both spatial constraints and speed demands. By scaling Δt, with their respective means, the ID accounts for individual variability in task engagement and reaction capabilities.

3.2 Experimental Design

To empirically assess the efficacy of this ID construct, we designed an experiment featuring six distinct conditions. These conditions were derived by pairing three plate radii (small, medium, large) with two minimum ball velocities (high, low). A within-subjects design ensured that each participant experienced all six conditions, and to mitigate potential learning and sequence effects, counterbalancing was employed. Practice trials were provided before formal data collection to familiarize participants with the task dynamics.

3.3 Performance Measures

The primary dependent variable was the duration for which participants could sustain performance before failing (i.e., letting a ball fall off the plate). Additional measures included: **Reaction time T** to visual stimuli, **Self-assessed skill scores**, recorded via brief post-task questionnaires. Preliminary analyses revealed that self-assessed skill scores did not significantly predict performance outcomes. Consequently, these scores were excluded from subsequent modeling efforts.

3.4 Predictive Model

Building on the validated ID, we formulated a predictive model for task duration (y), expressed as:

$$y = (w_2 \boldsymbol{ID} + \frac{w_1}{\boldsymbol{ID}} + w_0) \times \frac{\boldsymbol{R}}{\boldsymbol{V}} \tag{2}$$

where w_0, w_1 and w_2 are model coefficients determined through regression, and $\frac{\mathrm{R}}{\mathrm{V}}$ incorporates the critical ratio of plate radius to ball velocity. This functional form captures both direct and inverse relationships between the ID and duration, reflecting how game difficulty and participant characteristics jointly influence performance.

3.5 Model Evaluation and Application

Model evaluation using Mean Squared Error (MSE = 7.2992) and coefficient of determination (R2 = 0.7657) demonstrated a robust fit, affirming that the ID effectively characterizes performance across varying difficulty levels. Subsequently, we employed the model to identify optimal game parameter configurations aimed at achieving a target duration of 20–30 seconds—an interval deemed both suitably challenging and engaging based on participant feedback.

For instance, a 20-second target duration corresponds to an ID of approximately 99.19 when the plate radius is 5 and the minimum velocity is 0.2291, whereas a 30-second target duration aligns with an ID of about 83.29 at a radius of approximately 6.4735 and a minimum velocity of 0.2. These findings underscore the utility of the ID-driven methodology for systematically calibrating game parameters, allowing researchers and designers to balance challenge and user engagement in real-time multitasking environments.

4 User Study

This user study examined how variations in ball velocity and plate radius affect human multitasking performance in the Dual-Task Balancing Game. Using a within-subject design, all participants completed every experimental configuration, and a counterbalancing approach controlled for order effects. Both quantitative metrics and qualitative feedback provided a comprehensive assessment of how task difficulty influences performance.

4.1 Participants

The study recruited ten participants through online announcements and personal contacts. Of these, seven were students enrolled in a human-interface media course, and three were acquaintances of the researchers, added to meet the experimental requirements. Participant ages ranged from 20 to 25 years (two aged 20, one aged 21, three aged 23, and four aged 24), with one female and nine males.

Participants rated their gaming proficiency on a Likert scale from 1 (beginner) to 5 (expert). Two participants self-assessed as level 2, three as level 3 (intermediate), three as level 4, and one as level 5 (expert); none identified as a beginner. Additionally, a reaction-time test—requiring participants to click a circle promptly after its color changed—served as an objective measure of their reflexive abilities and multitasking aptitude.

4.2 Experimental Design

This investigation employed a within-subject design, ensuring each participant was exposed to all six task configurations. The configurations were established by combining two velocity conditions (0.3 and 0.5) with three plate radii (3, 5, and 7), producing the following pairs: (0.3, 3), (0.3, 5), (0.3, 7), (0.5, 3), (0.5, 5), and (0.5, 7).

To address potential order effects, the sequence of configurations was counterbalanced via a balanced Latin-square method. Participants also completed one-minute practice sessions before the actual data collection, ensuring that observed performance differences were attributable to the experimental conditions rather than variations in familiarity or skill transfer.

4.3 Procedure and Setup

Participants performed the **3D Dual-Task Balancing Game**, a scenario requiring them to balance a ball on two separate plates by tilting each plate using keyboard arrow keys. The difficulty was modulated by ball velocity (0.3 or 0.5) and plate radius (3, 5, or 7), where higher velocity demanded quicker responses, and smaller radii required greater precision.

The experiment took place in a controlled laboratory setting. Two PCs were arranged side-by-side, and participants alternated between them. While one PC was used for gameplay, the researcher prepared the subsequent trial on the other, ensuring a continuous, interruption-free experience. All environmental conditions were standardized, including lighting, seating arrangement, and the distance from the monitor.

Each configuration was tested in three trials, with the task continuing until the ball fell off the plate. After three trials, the configuration concluded. Participants completed all six configurations in an order determined by the balanced Latin-square arrangement.

4.4 Data Collection

The primary performance metric was **task completion time**, defined as the duration participants sustained the ball on the plate before it fell. This measure was recorded for each trial, enabling a quantitative assessment of how velocity and radius influenced task performance.

4.5 Metrics

In addition to quantitative data, qualitative feedback was collected post-task. Participants were asked about their overall impressions of the game and which configuration they found most enjoyable. All participants reported that the game was challenging, and most expressed a preference for the larger plate configurations, with no explicit preferences regarding velocity.

4.6 Ethical Considerations

All participants provided **informed consent** after receiving a clear explanation of the study's objectives and procedures. Data were anonymized and used exclusively for research purposes, and personal information was securely stored until the conclusion of the course, after which it was permanently deleted.

Table 1: Dataset

Velocity	Radius	Mean	Median	Std
0.3	3	7.041	6.775	1.363
0.3	5	12.923	12.760	2.718
0.3	7	19.341	19.620	4.445
0.5	3	5.453	5.500	0.450
0.5	5	8.204	8.275	1.076
0.5	7	14.271	12.490	4.157

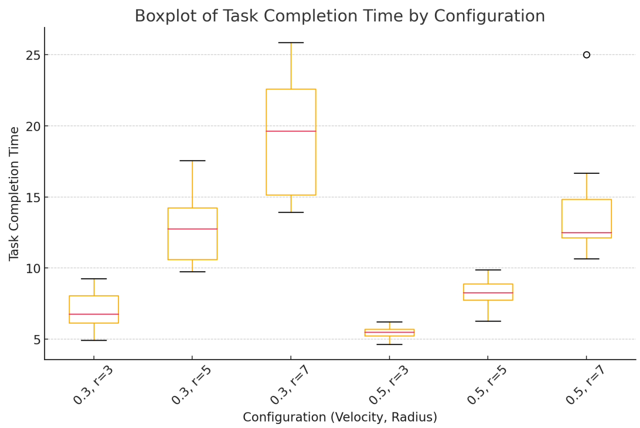

Figure 2: Task Completion Time

Table 2: Two-Way Repeated Measure ANOVA

Source	SS	DF	MS	F-Value	P-Value
Velocity(A)	215.616	1	215.616	32.757	< 0.001
Radius(B)	1127.228	2	563.614	46.166	< 0.001
AxB Interaction	36.752	2	18.376	4.544	0.025

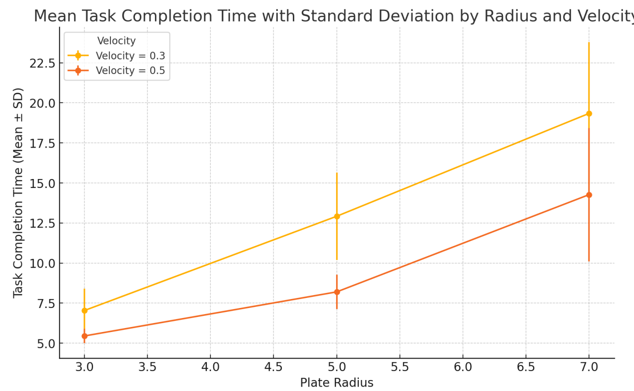

Figure 3: The interaction effect, showing that the difference in task completion times between the two velocities becomes more pronounced as the radius increases.

Although the study was not formally reviewed by an institutional ethics board due to its limited scope, all procedures followed standard ethical research practices involving human participants.

5 Results

This section presents the outcomes of the statistical analyses, highlighting the influence of ball velocity and plate radius on task completion time. A Two-Way Repeated Measures ANOVA examined the main effects and interaction effects of these factors. Descriptive and inferential results are provided, complemented by visualizations that illuminate performance trends across conditions.

5.1 Dataset

The dataset comprised six conditions formed by two velocity levels (0.3, 0.5) and three radius levels (3, 5, 7). Table 1 reports means and standard deviations for each condition, and Figure 3 illustrates the distribution of completion times.

5.2 Two-Way Repeated Measures ANOVA Test

A Two-Way Repeated Measures ANOVA was conducted to evaluate the effects of velocity and plate radius on task completion time (Table 2). Ten subjects completed all six velocity-radius combinations.

5.2.1 ***Result.*** The grand mean completion time across all conditions was approximately 11.2 seconds. Velocity had a significant effect: at 0.3, the mean time was about 13.1 seconds, whereas at 0.5, it was about 9.3 seconds ($F_{1,9} = 32.76$, $p < 0.001$).Radius also significantly affected completion times: a 3 radius yielded approximately 6.3 seconds, 5 about 10.6 seconds, and 7 about 16.8 seconds ($F_{2,18} = 46.17$, $p < 0.001$).

Figure 3 illustrates these trends, showing that as the plate radius increased, task difficulty also intensified. Moreover, the Velocity × Radius interaction was statistically significant ($F_{2,18} = 4.54$, $p = 0.025$).While differences between velocities were minimal at the smallest radius, this disparity grew as the radius increased. Post hoc tests (e.g., Scheffé) confirmed that the pronounced interaction was due chiefly to the increasingly larger performance gap at higher radii.

5.2.2 ***Summary.*** Both higher velocity and smaller radius improved task completion times. The interaction between these factors demonstrates that increasing velocity can more effectively counteract the challenges posed by larger radii. These findings underscore the importance of considering multiple factors concurrently when designing systems and interfaces to support human multitasking in complex environments.

6 Conclusion

The Dual-Task Balancing Game served as the principal platform for optimizing task difficulty. By integrating the proposed Index of Difficulty (ID), we developed a predictive model that achieved an MSE of 7.2992 and an R2-score of 0.7657. Although a comprehensive qualitative investigation was not undertaken, determining an optimal target duration remained challenging. Nevertheless, a subsequent user study revealed that most participants perceived the task as excessively difficult, preferring the widest disc configuration.

In light of these observations, the target duration was selected to align with participants' reported comfort levels. Additionally, by applying the model's recommended parameter settings, we could reliably achieve performance durations within the 20–30 second range. These findings underscore the efficacy of the predictive model in identifying and refining the game's difficulty, thus facilitating more adaptable and user-centered gameplay experiences.

Acknowledgments

This research builds on recent work in task switching [1].

References

[1] Alexander Lingler, Dinara Talypova, Jussi PP Jokinen, Antti Oulasvirta, and Philipp Wintersberger. 2024. Supporting Task Switching with Reinforcement Learning. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.

Received 09 December 2024; revised 11 December 2024; accepted 22 December 2024