

Assessing executive functions in the real world is virtually possible

Francisco Rocabado

Universidad Nebrija

Jon Andoni Duñabeitia (✉ jdunabeitia@nebrija.es)

Universidad Nebrija

Article

Keywords:

Posted Date: June 22nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1756406/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Executive functions have become a popular research topic since the beginning of the century. Their central role in our daily routines makes them the most representative form of behaviour regulation. Even though there is no general agreement about how they can be classified, there is common consensus that cognitive scientists need to understand their nature and how they interact to modulate human behaviour. Inhibitory control is one of the main exponents of executive functions, and the last decades have witnessed an exponential increase of studies on the inhibitory component thanks to the development of chronometric tasks associated with paradigms that allowed to explore how humans behave in contexts in which resistance to interference is needed. Among the different paradigms typically used, the Simon and flanker tasks are probably the most popular ones. Since their creation, many variants of these tasks have been developed in order to assess inhibitory control from different perspectives (e.g., in different samples or in combination with different research techniques). However, the use of these tasks has been typically governed by classical presentation modalities within laboratory settings. The accessibility of virtual reality technology has opened new research avenues to explore inhibitory skills using adaptations of classic paradigms such as the Simon and the flanker tasks to real-life contexts or settings. Thanks to virtual reality, research on executive function assessment can be undertaken with increased ecological validity, still ensuring laboratory-like controlled conditions and high precision in the measurements. Relying on this technology, here we present two state-of-the-art adaptations of the classical Simon and flanker tasks reinterpreted and adapted to real-world circumstances. Our results show that virtual reality is a reliable tool to assess inhibitory control that could provide valid experimental data with a high level of real-world transferability and generalizability.

Introduction

Most daily human interactions with the environment, such as walking in a crowded space, visiting a new city, or running errands, require a suitable response for achieving the goal. Here, Executive Functions (EF) play the main role in providing adequate responses in those contexts in which automatic or instinctive responses are not enough^[1], whether we find ourselves in a familiar situation or not. Those responses come from a group of top-down cognitive processes that allow us to go throughout our daily routine efficiently. In the long run, EF have been proven to have an important role in cognitive and psychosocial development^[2]. Furthermore, they have been pinned as one of the main predictors of students' academic performance and career trajectory development^[3-5] and are hence an important aspect of our physical and mental health.

Although there is no strict criterion for classifying EF, there is general agreement supporting a hierarchical framework that splits these abilities into three different core EF: Inhibition, working memory, and cognitive flexibility^[6-8]. From these core EF, higher-order EF abilities such as reasoning, problem-solving, and planning are distilled^[6]. In this sense, even though all EF play a significant role in our life successes, their acquirement and trajectories seem to be linked to the needs of the specific developmental stages^[9]. For

instance, it is not surprising that core EF fulfil their development early in our lifecycle before adolescence, whilst higher-order EF continue their development through adolescence, reaching ceiling capacity in early adulthood^[10,11].

Inhibitory control has been considered the most important foundational unit for EF^[7]. However, its nature and value as a foundational unit common within other EF is controversial to say the least since such a relevant role has been questioned more recently ^[8,12,13]. Nevertheless, regardless of the controversy, there is no doubt about its value in our daily lives. Frequently, the characteristics of the environment create situations in which salient (bottom-up) stimuli will involuntarily attract our attention^[14,15]. However, EF are typically capable of top-down distributing our attentional resources in a goal-oriented way so that we can selectively focus our attention while ignoring other interfering stimuli^[16]. In this way, and empowered by EF, our behavioural skills create the possibility of change and choice that help us to willingly regulate emotional, cognitive and behavioural responses without being under the control of environmental requirements, even when the circumstances might predispose another strong automatic response^[6,13,17]. Thus, it is important for goal-oriented achievements that come from making the right choices whilst avoiding temptation by regulating impulsive decisions and behaviours^[18,19].

In the last 50 years, cognitive psychologists have been vested in elaborating controlled laboratory tasks to measure inhibitory control performance. So far, nonverbal interference tasks have taken most of the research attention, and among these, the standouts are the Simon and flanker tasks^[18]. In 1967 J. Richard Simon and Alan P. Rudell^[20] introduced the first version of the task now known as the Simon task where participants were given two simple instructions: press a given key (e.g., F) for stimulus type A and another key (e.g., J) for stimulus type B (see Fig. 1). One stimulus is displayed at a time, and, although its position is irrelevant to the task, participants' responses are faster and more accurate when both target stimulus and response locations match than when they do not. This phenomenon, called the Simon effect, evidences a facilitating natural tendency to respond to the stimuli's location intermixed with a difficulty of inhibiting distracting and incongruous spatial information^[20]. A few years later, in 1973, Barbara A. Eriksen and Charles W. Eriksen²¹ created the well-recognised flanker task to complete the scarce literature on the field of visual search of their time. As in the Simon task, the instructions to complete the task were simple: participants had to respond by pressing one of the two designated keys depending on the orientation of each target stimulus (e.g., press F for stimulus A and J for stimulus B; see Fig. 1). However, unlike the Simon task, the target stimulus is presented in the centre of the screen and accompanied by flanking congruent or incongruent stimuli. In the original task, both the target and flanking stimuli were letters; every letter-string trial could display a target letter flanked by the same letter or by a different distracting letter. They found that reaction times were related to the level of competition flanking stimuli induced over the target stimulus. Thus, a target stimulus flanked by the same stimulus would produce more rapid and accurate responses than a target stimulus accompanied by a competitor, creating a response competition that we now call the flanker effect^[22].

Since their creation, both the Simon and flanker tasks have evolved to adapt their applicability over different presentation modalities to reach a broad spectrum of samples^[23–27]. As a result of these modifications, inhibitory control data from these paradigms with stimuli presented in the visual modality^[28–31], the auditory modality^[32,33] and the tactile modality^[34] can now be easily obtained¹. This has allowed researchers to assess participants from different age groups^[23,35] and special populations with and without medical conditions^[36,37].

For many years, these two tasks have helped capture individuals' inhibitory control by collecting data from the two critical conditions' trials (i.e., congruent and incongruent). Then, by contrasting the underlying processing differences by analysing reaction times and accuracy scores, an interference score (which is an indicative measure of a participants' inhibitory control capacity) can be established. While the resulting scores from these two tasks cannot be due to the same inhibitory control mechanisms involved during the task execution given that they typically show low correlation coefficients^[18], there's general agreement that both tasks tap into conflict monitoring and management. The distinction between them has been established on the basis of the type of conflict they generate and the hypothetical inhibition factor these conflicts represent – inhibition of prepotent responses in the case of the Simon task and resistance to distractor interference in the case of the flanker task^[38].

Aside from the taxonomic differences and the specific attentional control demands in the two tasks, the debate about their capacity to measure inhibitory control capacity in real-world activities remains open^[18]. Not surprisingly, these tasks have been used almost exclusively in laboratory settings, and this could obviously yield laboratory- and task-specific effects^[39]. The difficulty of translating laboratory-studied paradigms into a more ecological environment without sacrificing control over extraneous variables is a generalised conundrum for cognitive scientists, and improved evaluation strategies are still to be developed to capture the underlying mechanisms of inhibitory control capacity in real-world activities with high precision and ecological validity.

The technological revolution experienced in the last 30 years made it possible to create Virtual Reality (VR) systems that are now well implemented in different research fields, such as neuroscience and psychology, as a reliable assessment tool^[40]. New generation VR headsets are price-accessible and powerful enough to provide high-resolution stimuli presentation within realistic and interactive 3D environments, providing increased ecological validity, flexibility, sensory feedback, and performance recording^[40,41] – see Gaggioli^[42] for an early review of VR application. In fact, there are many successful examples of the use of VR as an assessment tool for skills related to EF, such as working memory and task-switching^[43], cognitive load in navigation^[44], memory^[45], behavioural responses in eating disorders^[46] and, more recently, cognitive function in terms of reasoning and attention switching^[47].

Likewise, a few examples can be found regarding inhibitory control assessment in VR. In 2018, Olk et al. successfully adapted the flanker task and increased its ecological validity by using daily objects instead of letters or arrows as stimuli. Later in 2019, Roberts et al. compared responses of the arrow version of

the flanker task from a real laboratory setting and a matching replica scenario of the same laboratory in VR, showing that assessment from both scenarios is comparable and proves the utility of VR as an assessment tool. The most recent adaptation of the flanker task consists of projected typographic symbols over panels across a street^[48]. In the same line, Gupta and colleagues^[49] translated the Simon task to a VR setting, aiming to assess alternative cues, response modalities and VR ergonomics by projecting 2D stimuli through the VR headset. As seen, both tasks have been successfully implemented in VR settings, essentially mimicking 2D scenarios in a 3D context.

To take full advantage of VR technology seeking to overcome current flanker and Simon tasks VR-adaptation limitations, here we propose a new strategy to assess inhibitory control by translating the classical tasks to a more realistic type of context, including humans as target stimuli and their behaviour as the critical conditions. Here, we aim to explore this technology's advantages as an assessment tool for inhibitory control in a close-to-reality environment without compromising assessment reliability by introducing a new user-friendly adaptation of both tasks within a complex real-world scenario that closely resembles a school classroom. Two experimental tasks were designed for each paradigm (Simon and flanker), one in 2D format and one being a 3D-adapted version of the same task.

Experiment 1: Simon Task

Participants

We recruited 36 participants from the Universitat de València with ages between 18 to 32 years old (mean = 22.84, SD = 3.90; 14 females) and normal or corrected-to-normal vision. All participants were informed that data would be collected anonymously during the session and signed an informed consent form before starting. All experimental protocols were approved by the Ethics Board of the Universidad Nebrija (Comité de Ética en Investigación – CEI), and the studies were carried out in accordance with the general guidelines of the Ethics Committee.

Materials and Procedure

2D Simon task

All stimuli were presented on a 15.6" laptop screen from a distance of about 55cm. The stimuli were one blue and one red colour solid square displayed on a white background. Each trial began with a central fixation cross (+) for 500ms. Participants were asked to press the "F" key if they saw a red square and the "J" Key if they saw a blue square. They were asked to respond as fast as possible, avoiding making mistakes. A maximum response time of 3000ms was accepted.

All trials pertained to three critical experimental conditions (congruent, incongruent, and neutral). Trials were defined as congruent when the target was displayed on the same side as the correct response key (e.g., a red square positioned on the left side of the screen). Incongruent trials were those on which the target location and the correct response were on opposite sides (e.g., a red squared displayed on the right

side of the screen). Finally, when the target was displayed centred to the fixation cross, these trials were defined as neutral (see Fig. 1 for illustration of the stimulus type employed in our task).

The 2D experimental task was built using Gorilla Experiment Builder^[50] and executed within the same web-based platform (www.gorilla.sc).

Table 1. Mean accuracy and latency results per condition, modality, and task. The standard deviation is presented in parentheses.

Task	Stimulus Type	Accuracy	RT (in ms)
Simon 2D	Congruent	0.96 (.04)	420 (75.53)
	Incongruent	0.89 (.08)	462 (87.68)
	Neutral	0.95 (.04)	419 (67.53)
Simon 3D	Incongruent	0.98 (.02)	641 (110.06)
	Congruent	0.91 (.06)	679 (100.75)
	Neutral	0.96 (.03)	636 (102.64)
Flanker 2D	Congruent	0.98 (.02)	441 (71.90)
	Incongruent	0.90 (.14)	515 (123.58)
	Neutral	0.98 (.02)	441 (69.60)
Flanker 3D	Congruent	0.98 (.02)	372 (82.40)
	Incongruent	0.94 (.06)	422 (89.23)
	Neutral	0.98 (.03)	450 (95.96)

Simon 3D task

The chosen stimuli consisted of three male human avatars that were programmed to perform two different actions: clapping hands and raising a hand. Each action was displayed within a virtual reality environment resembling a classroom scenario. One of the avatars was positioned in front of the participant's point of view and centred within the scenario, and the other two were placed located at his left and right sides within the field of view. All three avatars were displayed simultaneously and programmed to remain in an idle position. Each trial began with a neutral auditory stimulus during 500ms that alerted participants of the beginning of a new trial. After 500ms, one of the avatars would perform one of the programmed movements. As in the 2D version, a maximum response time of 3000ms was permitted (see Fig. 2). A video of the virtual environment and the task is provided at <https://doi.org/10.6084/m9.figshare.19984631>.

Within the virtual environment, each controller became a virtual hand; thus, participants held each virtual hand with the corresponding real hand and were asked to give quick and accurate responses by pulling the trigger in each controller to indicate if the avatar performing a movement was either clapping (left trigger) or raising a hand (right trigger). Similar to the 2D version of the task, all trials pertained to one out of three critical experimental conditions (congruent, incongruent, and neutral). Trials were defined as congruent when the moving avatars' action was performed on the same side as the correct response trigger (e.g., leftmost avatar clapping). Similarly, incongruent trials were defined as those in which the critical avatar's location and the correct response trigger were on opposite sides (e.g., the rightmost avatar clapping). Finally, when the moving avatar was the one centred to the participants' point of view, these trials were defined as neutral.

The 3D experimental task was created using Vizard 6.0 (WorldViz), a Python-based software (Python v. 2.7.12). The experiment script was executed on a high-end gaming laptop (MSI GL76) computer equipped with an Intel Core i7-10750H (2.6 Hz), running Windows 10 operating system (64 bit), 32 GB RAM, and an NVIDIA GeForce RTX 2070 video card. To ensure and maintain high-performance connections between the PC and the VR HMDs, battery-saving settings were disabled. 3D stimuli were presented through the HTC Vive Pro HMD (HTC Vive Pro, 2018) at 2880 × 1600-pixel resolution (1440 × 1600 per eye) and 90-Hz refresh rate. Thus, providing 110 degrees of field of view and high immersive experience, made of a high-quality display and a stable tracking system^[51].

Regardless of the task version (2D or 3D), all began with a practice period in order to familiarise participants with the task. This practice included 12 trials, four from each condition. After the practice, the experimental trials followed, including 48 trials per condition. Experimental trials were distributed across three blocks. Each block included 16 trials per condition that were randomly presented. Overall, the 2D task was completed in around 5 minutes and the 3D task in about 8 minutes. Between the two versions of the task a 15-minute distracting task was presented. The presentation of the two task modalities was counterbalanced across participants.

Results

Collected data were processed and cleaned in RStudio^[52] and analysed with JASP^[53]. Descriptive analyses were undertaken to ascertain reaction times and accuracy (see Table 1). Mean reaction times (RT) were computed for each condition and participant at a trial level by including only accurate responses. Additionally, participants' RT that were 2.5 SD faster or slower than the mean RT per condition or those associated with timed-out responses were rejected (2.94% of the data in the 2D modality and 1.86% of the data in the 3D version).

We carried out a 3 (Stimulus Type: congruent, incongruent, and neutral) × 2 (Task Modality: 2D and 3D) repeated-measures ANOVA on the RT data. Significant main effect of Stimulus Type was found ($F(2, 70) = 62.016, p < 0.001, \eta_p^2 = 0.639$). Post hoc analyses revealed that differences occurred between congruent and incongruent conditions ($M_{Diff} = -39.952, SE = 4.326, p_{bonf} < 0.001$), reflecting the classical Simon

interference effect, and between incongruent and neutral conditions ($M_{Diff} = 43.296$, $SE = 4.326$, $p_{bonf} < 0.001$), showing an incongruency effect. No significant difference was found between congruent and neutral conditions ($M_{Diff} = 3.343$, $SE = .773$, $p_{bonf} = 1.000$). Additionally, the main effect of Task Modality was also significant ($F(1, 35) = 189.223$, $p < 0.001$, $\eta_p^2 = 0.844$), being latencies from the 3D modality larger than the 2D ($M_{Diff} = 218.470$). Importantly, there was not an interaction between Stimulus Type and Task Modality ($F(2, 35) = 0.271$, $p = 0.277$, $\eta_p^2 = 0.008$) (see Fig. 3).

A similar repeated-measures ANOVA was performed on the accuracy scores from both tasks. When sphericity assumptions were violated, the Greenhouse-Geisser correction was applied. A significant main effect of Stimulus Type was found ($F(1.652, 70) = 48.783$, $p < 0.001$, $\eta_p^2 = .582$). Post hoc analyses showed significant differences between congruent and incongruent conditions ($M_{Diff} = 0.072$, $SE = 0.006$, $p_{bonf} < 0.001$; namely, a Simon effect), and between incongruent and neutral conditions ($M_{Diff} = -0.056$, $SE = .006$, $p_{bonf} < 0.001$; namely, an incongruence effect). No significant difference was found between congruent and neutral conditions ($M_{Diff} = 0.016$, $SE = 0.008$, $p_{bonf} = 0.110$). The main effect of Task Modality was also significant ($F(1, 35) = 6.535$, $p = 0.015$, $\eta_p^2 = 0.157$), being responses on the 3D modality more accurate than in the 2D modality ($M_{Diff} = 0.016$). Finally, there was not an interaction between Stimulus Type and Task Modality ($F(1.458, 35) = 0.225$, $p = 0.728$, $\eta_p^2 = 0.006$) (see Fig. 3).

Discussion

The 2D and 3D versions of the Simon task showed a markedly similar response and accuracy pattern across all stimulus type conditions. In both settings, incongruent stimuli elicited longer response latencies compared to congruent and neutral stimuli, and classic Simon effects and incongruency effects were replicated both in the 2D and 3D versions of the paradigm. RT significantly differed between task modalities, with participants showing shorter response latencies when the task was performed in the 2D context, an effect that can be easily explained by core differences in the stimuli presentation. Whereas in the 2D version of the task the stimuli were displayed without a movement component, in the 3D version participants had to hold their response until the avatar movement was evident. This interpretation aligns with the small difference between task modalities found in terms of accuracy (which should not be influenced by the time needed by the characters to initiate a movement), and with the similar pattern of effects found across modalities occur in all task conditions. As a conclusion, our results showed that both the 2D and 3D tasks are equally capable of capturing participants' inhibitory control towards prepotent responses as measured by the Simon interference effect.

Experiment 2: Flanker Task

Participants

For this experimental session, data from 46 participants from Nebrija University with ages between 18 to 52 years old ($M_{age} = 25.23$, $SD = 7.05$; 31 females) and normal or corrected-to-normal vision were

collected. Before starting the experimental session, participants were informed that data would be collected anonymously, and they signed an informed consent.

Materials and Procedure

Flanker 2D task

All stimuli were presented on a 15.6" laptop set at a distance of around 50cm to the participants' eyes. A series of arrays of five arrows were employed as experimental stimuli, including one central target arrow, and two flanking arrows at each side. Each trial began with a central fixation cross (+) for 500ms, followed by one stimuli array. The target stimulus was displayed for 3000ms or until a response was given. If a response was given, this was followed by a between-trial blank space lasting for 500ms. All participants were encouraged to keep their index fingers always positioned over the "F" key for the left hand and "J" key for the right one on the keyboard and they were prompted to respond as fast as possible while avoiding making mistakes. Participants' task was to press the button corresponding to the direction of the central target arrow: they had to press "F" if the target stimulus pointed to the left and "J" when the target pointed to the right. Experimental trial conditions were defined as follows: congruent when all five arrows pointed towards the same direction, incongruent when the central and the flanking arrows pointed in the opposite direction, and neutral when dashes instead of arrows flanked the target arrow.

As in the case of the Simon task, the 2D version of the flanker task was built with Gorilla Experiment Builder^[50].

Flanker 3D task

Primary experimental stimuli consisted of five male human avatars programmed to spin to the left or the right. The programmed movement took place within 400ms, and once the movement was completed, the avatar would return to its initial position. All five avatars were presented within a scenario that resembled a classroom. Following the same vein as in the classical flanker task, one of the avatars was positioned in the centre of the scene, in front of the participant's point of view, and flanked by two avatars on each side. All avatars were displayed simultaneously in a synchronised idle position, separated by the same distance. In order to maximise the presence of interference effects and after a series of pilot tests, the target avatar began to spin 200ms after the flanker avatars did. Each trial began with a neutral auditory stimulus that lasted 500ms, alerting participants of the beginning of a new trial. To avoid stimuli overlapping, the experimental trial would begin 500ms after the sound. Participants were asked to respond using the provided controllers by paying attention to the spinning direction of the central avatar: they were asked to pull the trigger of the left-hand controller if the target avatar spined to the left, and to pull the right-hand controller if he spun to the right. Experimental trial conditions were defined as congruent when all five avatars spun in the same direction, incongruent when the central avatar and the flanker avatars spun in opposite directions, and neutral when the central avatar performed his movement while the flanker avatars maintained their idle position. An illustrative schematic example is presented in

Fig. 4 and a video of the virtual environment, and the task is provided at <https://doi.org/10.6084/m9.figshare.19984631>.

The 3D version of the flanker task was created using Vizard 6.0 (WorldViz), and stimuli were presented through the HTC Vive Pro HMD (HTC Vive Pro, 2018), mimicking the 3D data collection associated with the Simon task reported in Experiment 1.

In both 2D and 3D versions, experimental trials were distributed across three blocks, each including 16 trials per condition randomly presented. Both versions began with 12 practice trials (four trials per condition). Overall, the 2D task lasted between 5 to 7 minutes whereas the 3D task was completed between 8 to 11 minutes. A distractor task was placed between our two presentation modalities. The distraction task lasted 15 minutes. All experimental tasks were presented in a counterbalanced order between participants.

Results

All collected data were processed and cleaned in RStudio^[52] and analysed in JASP^[53]. Descriptive analyses were performed to ascertain RT and accuracy (see Table 1). Mean RT were computed for each condition and participant at a trial level, including only accurate responses. Moreover, participants' RT that were 2.5 SD faster or slower than the mean or those associated with timed-out responses were trimmed from any analysis (2.78 % and 2.48% of trimmed data from the 2D and the 3D task modalities, respectively).

We first carried out a 3 (Stimulus Type: congruent, incongruent, and neutral) x 2 (Task Modality: 2D and 3D) repeated-measures ANOVA on the RT data. When sphericity assumptions were violated, Greenhouse-Geisser correction was applied. A significant main effect of Stimulus Type was found ($F(1.354, 60.923) = 67.938, p < 0.001, \eta_p^2 = 0.602$). Post hoc analyses showed significant differences between congruent and incongruent conditions (namely, a flanker effect; $M_{Diff} = -61.910, SE = 5.370, p_{bonf} < 0.001$), incongruent and neutral conditions (an incongruity effect; $M_{Diff} = 22.924, SE = 5.370, p_{bonf} = 0.001$) and between congruent and neutral condition (a congruency effect; $M_{Diff} = -38.987, SE = 5.370, p_{bonf} < 0.001$). Additionally, a significant main effect of Task Modality was found ($F(1, 45) = 30.081, p < 0.001, \eta_p^2 = 0.401$), being the latencies for the 2D modality longer in comparison to the 3D ($M_{Diff} = 51.205$). Finally, the interaction between Stimulus Type and Task Modality was also significant ($F(1.317, 59.282) = 41.683, p < 0.001, \eta_p^2 = 0.481$). Post hoc analyses showed that latencies were significantly larger in the 2D than in the 3D modality for congruent and incongruent conditions ($p_{bonf} < 0.001$), but no statistical differences were found between task modalities for the neutral conditions ($p_{bonf} = 1.00$; see Fig. 5).

For accuracy scores, a similar 3x2 repeated measures ANOVA was performed. A significant main effect of Stimulus Type was found ($F(1.074, 48.341) = 28.740, p < 0.001, \eta_p^2 = 0.390$). Post hoc analyses showed significant statistical differences between congruent and incongruent conditions (namely, a flanker effect;

$M_{\text{Diff}} = 0.063$, $SE = 0.009$, $p_{\text{bonf}} < .001$) and between incongruent and neutral conditions (an incongruity effect; $M_{\text{Diff}} = -0.059$, $SE = 0.009$, $p_{\text{bonf}} < 0.001$); on the contrary, no significant difference was found between the congruent and neutral conditions ($M_{\text{Diff}} = 0.004$, $SE = 0.009$, $p_{\text{bonf}} = .697$). No significant main effect on Task Modality was found ($F(1, 45) = 1.524$, $p = 0.223$, $\eta_p^2 = 0.033$). Finally, no interaction was found ($F(1.079, 48.564) = 1.470$, $p = 0.235$, $\eta_p^2 = 0.032$) (see Fig.5).

Discussion

The classical flanker effect was replicated in the 2D task modality, in which incongruent trials generated longer latencies and higher error rates compared to both congruent and neutral conditions. Importantly, the 3D modality of the flanker task yielded strikingly similar results as well: incongruent trials generated longer latencies and larger error rates as compared to congruent trials, showing an instance of the well-known flanker effect with a dramatised VR version of the task.

Unlike the 3D Simon task used in Experiment 1, in Experiment 2 the sequence of programmed movements was not complex, which could explain the shorter latencies observed in the VR modality as compared to the 2D version, especially in the congruent and incongruent conditions. An effect that deserves a mention is the relatively high RT associated with the items from the neutral condition in the 3D version of the task (450ms vs. 422ms in the incongruent and 372ms in the congruent condition, respectively). This effect can be understood in the specific context of the 3D task we devised, in which participants first previewed the beginning of the movement of the flanking characters 200ms before the target (central) character started to spin. In neutral trials, the flanking characters did not spin, and participants were most likely withholding response being conditioned by the expectancy of a trial beginning with the flanking avatars' movement. Thus, it is reasonable to believe that participants would carefully await during a neutral trial before giving any response. Hence, we interpret this condition-specific effect as a task-induced result that should not deviate the focus on the impact of the findings. Inhibitory control assessment by the flanker task has typically relied on the differential score between incongruent and congruent conditions, and here we showed that such conditions presented in a 3D realistic variant can perfectly capture participants' resistance to distractor interference the same way the classical 2D version of the task does.

General discussion

Virtual reality (VR) technology has been increasingly popular in recent years among healthcare providers, and its use for the assessment and training of cognitive skills has recently received a great deal of attention. In essence, VR allows users to interact with 3D environments using multiple senses, and in the current study we demonstrate the usefulness of employing a head-mounted device (HMD) with a three-dimensional environment to generate a completely immersive experience in which specific aspects of inhibitory control can be correctly measured.

Regardless of the procedure^[47,54] or the aim of the study^[48,49,55], none of the previous VR adaptations of the Simon and flanker tasks were capable of making the most of the possibilities offered by virtual reality

when transferring and testing these tasks in a creative, plausible and realistic way. The current work represents a new ecologically valid alternative for assessing inhibitory control as a relevant construct of executive functions. We present two state-of-the-art variants of the well-known Simon and flanker tasks in a fully immersive quasi-realistic virtual environment that aims to translate classical laboratory paradigms to real-world situations.

Through two experimental studies we have presented evidence showing that it is possible to capture a snapshot of certain cognitive processes as measured by reaction times and response accuracy by presenting stimuli in a context close to reality using VR. Stoffels and van der Molen's (1988) arrow representation of the flanker task used material that implied orienting directionality to generate the interference score between congruent and incongruent directions. Based on this premise, our task aimed to generate the same conflict by relying on the orienting movement made by the avatars. This approach enables us to generate a more natural expression of the task that allows us to generate a more naturalistic assessment of inhibitory control that, as a result, helps us deepen a concept that, until now, has been limited to 2D scenarios. In this sense, our results show that both 2D and 3D presentation modalities are equally capable of capturing classic indices related to resistance to interference and inhibitory control. Our results open the path toward new possibilities for the real-world assessment of human inhibitory skills. We are confident that this study will open doors to future adaptations of classical cognitive tasks to more naturalistic scenarios.

Although there is still a long way to go in the process of standardising the use of VR systems as tools to deliver psychometric tests, the use of techniques that allow us to evaluate cognitive processes using chronometric-based behavioural responses with high levels of ecological validity will positively impact the quality of the data associated with the measurements. Since the 2020 announcement of Facebook's intentions to move its platform into the metaverse, many R&D companies have started to allocate economic and technological resources to the advancement of the virtual universe, boosting the expansion of VR systems. In accordance with these changes, we firmly believe that future research should make an effort to create and validate adaptations of the well-established paradigms from cognitive science for VR platforms.

The use of VR for cognitive assessment incurs some costs that might slow down the generalisation and globalisation of this technology. We believe that the scientific community should move towards an open-source practice where tools and methods are shared. In this vein, we aim to provide the community with the use of these two new adaptations of the flanker and Simon task with the hope that it will positively impact our understanding of EF assessment. To this end, our scripts have been made available at <https://doi.org/10.6084/m9.figshare.19984631>. While there are still some rough edges in terms of the immersiveness of the VR scenarios that will certainly experience important changes in the following years, it is important to restate that these do not seem to significantly alter the results from the tasks: inhibitory skills and incongruity effects can be effectively measured in close-to-real VR scenarios.

Declarations

Acknowledgements (not compulsory)

The authors are grateful to the Universitat de València which provided facilities to run the experiments. The authors thank the study participants, without whom this work would not have been possible; and to Jose Rocabado Rocha for providing assistance during the programming of the VR experiment.

Author contributions statement

F.R. and J.A.D. conceived and designed the experiment; F.R. programmed and conducted the experiments; F.R. and J.A.D. analysed the data and interpreted the data; F.R. drafted the manuscript; J.A.D. provided critical revisions of the manuscript.

Data availability statement

The data and the scripts can be accessed at <https://doi.org/10.6084/m9.figshare.19984631>

Ethics declarations

Competing interests

The authors declare no competing interests.

References

1. Burgess, P. & Simons, J. Theories of frontal lobe executive function: Clinical applications. *Eff. Rehabil. Cogn. Deficits* 211–231 (2005) doi:10.1093/acprof:oso/9780198526544.003.0018.
2. Gilbert, S. J. & Burgess, P. W. Executive function. *Curr. Biol.* **18**, R110–R114 (2008).
3. Munakata, Y. & Michaelson, L. E. Executive Functions in Social Context: Implications for Conceptualising, Measuring, and Supporting Developmental Trajectories. *Annu. Rev. Dev. Psychol.* **3**, null (2021).
4. Singer, B. D. & Bashir, A. S. What Are Executive Functions and Self-Regulation and What Do They Have to Do With Language-Learning Disorders? *Lang. Speech Hear. Serv. Sch.* **30**, 265–273 (1999).
5. Toll, S. W. M., Van der Ven, S. H. G., Kroesbergen, E. H. & Van Luit, J. E. H. Executive Functions as Predictors of Math Learning Disabilities. *J. Learn. Disabil.* **44**, 521–532 (2011).
6. Diamond, A. Executive Functions. *Annu. Rev. Psychol.* **64**, 135–168 (2013).
7. Miyake, A. *et al.* The unity and diversity of executive functions and their contributions to complex 'Frontal Lobe' tasks: a latent variable analysis. *Cognit. Psychol.* **41**, 49–100 (2000).
8. Friedman, N. P. & Miyake, A. Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex* **86**, 186–204 (2017).
9. Vestberg, T., Reinebo, G., Maurex, L., Ingvar, M. & Petrovic, P. Core executive functions are associated with success in young elite soccer players. *PLOS ONE* **12**, e0170845 (2017).

10. Best, J. R. & Miller, P. H. A Developmental Perspective on Executive Function: Development of Executive Functions. *Child Dev.***81**, 1641–1660 (2010).
11. Romine, C. B. & Reynolds, C. R. A Model of the Development of Frontal Lobe Functioning: Findings From a Meta-Analysis. *Appl. Neuropsychol.***12**, 190–201 (2005).
12. Miyake, A. & Friedman, N. P. The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Curr. Dir. Psychol. Sci.***21**, 8–14 (2012).
13. Munakata, Y. *et al.* A unified framework for inhibitory control. *Trends Cogn. Sci.***15**, 453–459 (2011).
14. Serences, J. T. *et al.* Coordination of Voluntary and Stimulus-Driven Attentional Control in Human Cortex. *Psychol. Sci.***16**, 114–122 (2005).
15. Wang, B. & Theeuwes, J. Salience determines attentional orienting in visual selection. *J. Exp. Psychol. Hum. Percept. Perform.***46**, (2020).
16. Theeuwes, J. Top-down and bottom-up control of visual selection. *Acta Psychol. (Amst.)***135**, 77–99 (2010).
17. Wessel, J. R. Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. *Psychophysiology***55**, e12871 (2018).
18. Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L. & Mikulinsky, R. Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cogn. Res. Princ. Implic.***5**, 7 (2020).
19. Banich, M. T. & Depue, B. E. Recent advances in understanding neural systems that support inhibitory control. *Curr. Opin. Behav. Sci.***1**, 17–22 (2015).
20. Simon, J. R. & Rudell, A. P. Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *J. Appl. Psychol.***51**, 300–304 (1967).
21. Eriksen, B. A. & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.***16**, 143–149 (1974).
22. Ridderinkhof, K. R., Wylie, S. A., van den Wildenberg, W. P. M., Bashore, T. R. & van der Molen, M. W. The arrow of time: Advancing insights into action control from the arrow version of the Eriksen flanker task. *Atten. Percept. Psychophys.***83**, 700–721 (2021).
23. Ambrosi, S., Śmigasiewicz, K., Burle, B. & Blaye, A. The dynamics of interference control across childhood and adolescence: Distribution analyses in three conflict tasks and ten age groups. *Dev. Psychol.***56**, 2262–2280 (2020).
24. Bidet-Ildei, C. & Bouquet, C. Motor Knowledge Modulates Attentional Processing during Action Judgment. *ATHENS J. Soc. Sci.***2**, 249–262 (2015).
25. Enns, J. T. & Akhtar, N. A Developmental Study of Filtering in Visual Attention. *Child Dev.***60**, 1188 (1989).
26. Kerzel, D., Weigelt, M. & Bosbach, S. Estimating the quantitative relation between incongruent information and response time. *Acta Psychol. (Amst.)***122**, 267–279 (2006).

27. Stoffels, E. J. & van der Molen, M. W. Effects of visual and auditory noise on visual choice reaction time in a continuous-flow paradigm. *Percept. Psychophys.***44**, 7–14 (1988).
28. McDermott, J. M., Pérez-Edgar, K. & Fox, N. A. Variations of the flanker paradigm: Assessing selective attention in young children. *Behav. Res. Methods***39**, 62–70 (2007).
29. Paap, K. R. & Greenberg, Z. I. There is no coherent evidence for a bilingual advantage in executive processing. *Cognit. Psychol.***66**, 232–258 (2013).
30. Waszak, F., Scientificque, D. L. R., Li, S. & Hommel, B. The development of attentional networks: crosssectional findings from a life span. (2010).
31. Zhong, Q., Proctor, R. W., Xiong, A. & Vu, K.-P. L. Transfer of incompatible spatial mapping to the vertical Simon task generalises across effectors but not stimulus features. *Atten. Percept. Psychophys.***82**, 1–11 (2020).
32. Ulrich, R., Prislán, L. & Miller, J. A bimodal extension of the Eriksen flanker task. *Atten. Percept. Psychophys.***83**, 790–799 (2021).
33. Vu, K.-P. L., Proctor, R. W. & Urcuioli, P. Transfer effects of incompatible location-relevant mappings on a subsequent visual or auditory Simon task. *Mem. Cognit.***31**, 1146–1152 (2003).
34. Baciero, A., Uribe, I. & Gomez, P. The tactile Eriksen flanker effect: A time course analysis. *Atten. Percept. Psychophys.***83**, 1424–1434 (2021).
35. Rueda, M. R. *et al.* Development of attentional networks in childhood. *Neuropsychologia***42**, 1029–1040 (2004).
36. Paap, K. R. & Sawi, O. The role of test-retest reliability in measuring individual and group differences in executive functioning. *J. Neurosci. Methods***274**, 81–93 (2016).
37. Wang, P. *et al.* Perceptual and response interference in Alzheimer's disease and mild cognitive impairment. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***124**, 2389–2396 (2013).
38. Kornblum, S., Hasbroucq, T. & Osman, A. Dimensional overlap: cognitive basis for stimulus-response compatibility—a model and taxonomy. *Psychol. Rev.***97**, (1990).
39. Winkler, R. L. & Murphy, A. H. Experiments in the laboratory and the real world. *Organ. Behav. Hum. Perform.***10**, 252–270 (1973).
40. Vasser, M. & Aru, J. Guidelines for immersive virtual reality in psychological research. *Curr. Opin. Psychol.***36**, 71–76 (2020).
41. de Gelder, B., Kätsyri, J. & de Borst, A. W. Virtual reality and the new psychophysics. *Br. J. Psychol.***109**, 421–426 (2018).
42. Gaggioli, A. Using virtual reality in experimental psychology. in *Towards cyberpsychology: Mind, cognition and society in the internet age* 157–174 (IOS Press, 2001).
43. Jansari, A. S., Froggatt, D., Edgington, T. & Dawkins, L. Investigating the impact of nicotine on executive functions using a novel virtual reality assessment. *Addiction***108**, 977–984 (2013).
44. Armougum, A., Orriols, E., Gaston-Bellegarde, A., Marle, C. J.-L. & Piolino, P. Virtual reality: A new method to investigate cognitive load during navigation. *J. Environ. Psychol.***65**, 101338 (2019).

45. Siemerkus, J., Irle, E., Schmidt-Samoa, C., Dechent, P. & Weniger, G. Egocentric spatial learning in schizophrenia investigated with functional magnetic resonance imaging. *NeuroImage Clin.***1**, 153–163 (2012).
46. Gorini, A., Griez, E., Petrova, A. & Riva, G. Assessment of the emotional responses produced by exposure to real food, virtual food and photographs of food in patients affected by eating disorders. *Ann. Gen. Psychiatry***9**, 30 (2010).
47. Williams, R. M., Alikhademi, K. & Gilbert, J. E. Design of a toolkit for real-time executive function assessment in custom-made virtual experiences and interventions. *Int. J. Hum.-Comput. Stud.***158**, 102734 (2022).
48. Ribeiro, N., Vigier, T. & Prié, Y. Tracking motor activity in Virtual Reality to reveal cognitive functioning: A preliminary study. *Int. J. Virtual Real.***21**, 30–46 (2021).
49. Gupta, A., Edwards III, H. M., Rodriguez, A. R., McKindles, R. J. & Stirling, L. A. Alternative cue and response modalities maintain the Simon effect but impact task performance. *Appl. Ergon.***100**, 103648 (2022).
50. Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. & Evershed, J. K. Gorilla in our midst: An online behavioral experiment builder. *Behav. Res. Methods***52**, 388–407 (2020).
51. Angelov, V., Petkov, E., Shipkovenski, G. & Kalushkov, T. Modern Virtual Reality Headsets. in *2020 International Congress on Human-Computer Interaction, Optimisation and Robotic Applications (HORA)* 1–5 (2020). doi:10.1109/HORA49412.2020.9152604.
52. RStudio Team. *RStudio: Integrated Development Environment for R*. (RStudio, PBC., 2020).
53. JASP Team. *JASP (Version 0.16.2)[Computer software]*. (2022).
54. Olk, B., Dinu, A., Zielinski, D. J. & Kopper, R. Measuring visual search and distraction in immersive virtual reality. *R. Soc. Open Sci.***5**, 172331.
55. Roberts, A. C. *et al.* Assessing the suitability of virtual reality for psychological testing. *Psychol. Assess.***31**, 318–328 (2019).

Figures

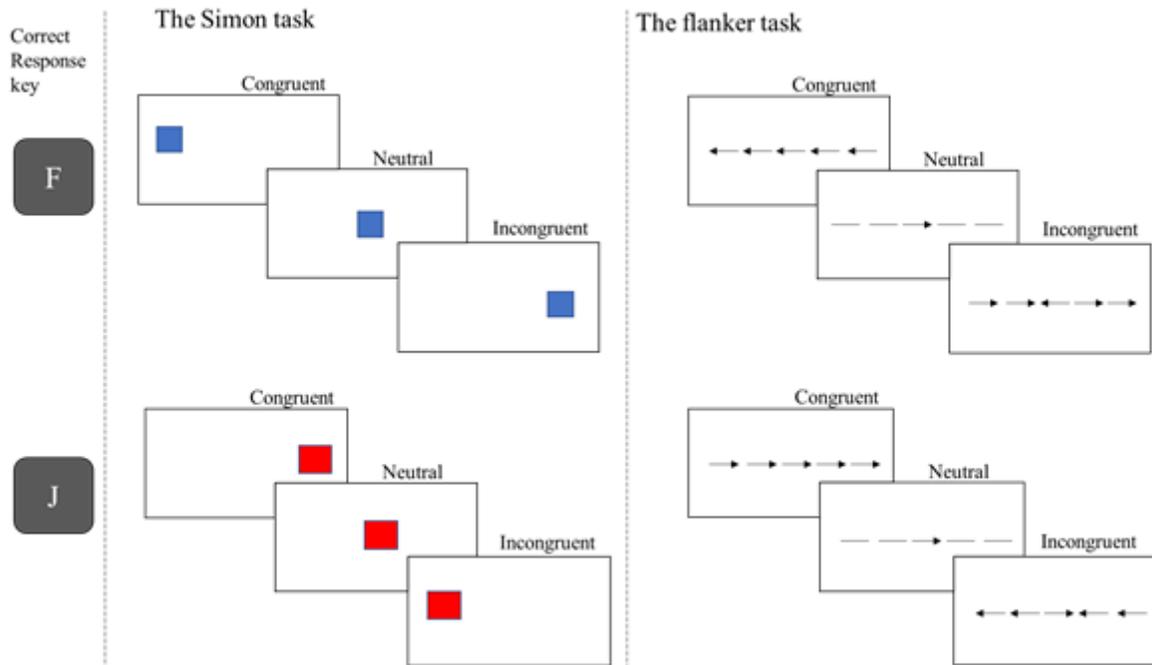


Figure 1

Typical trial structure The Simon and flanker task.

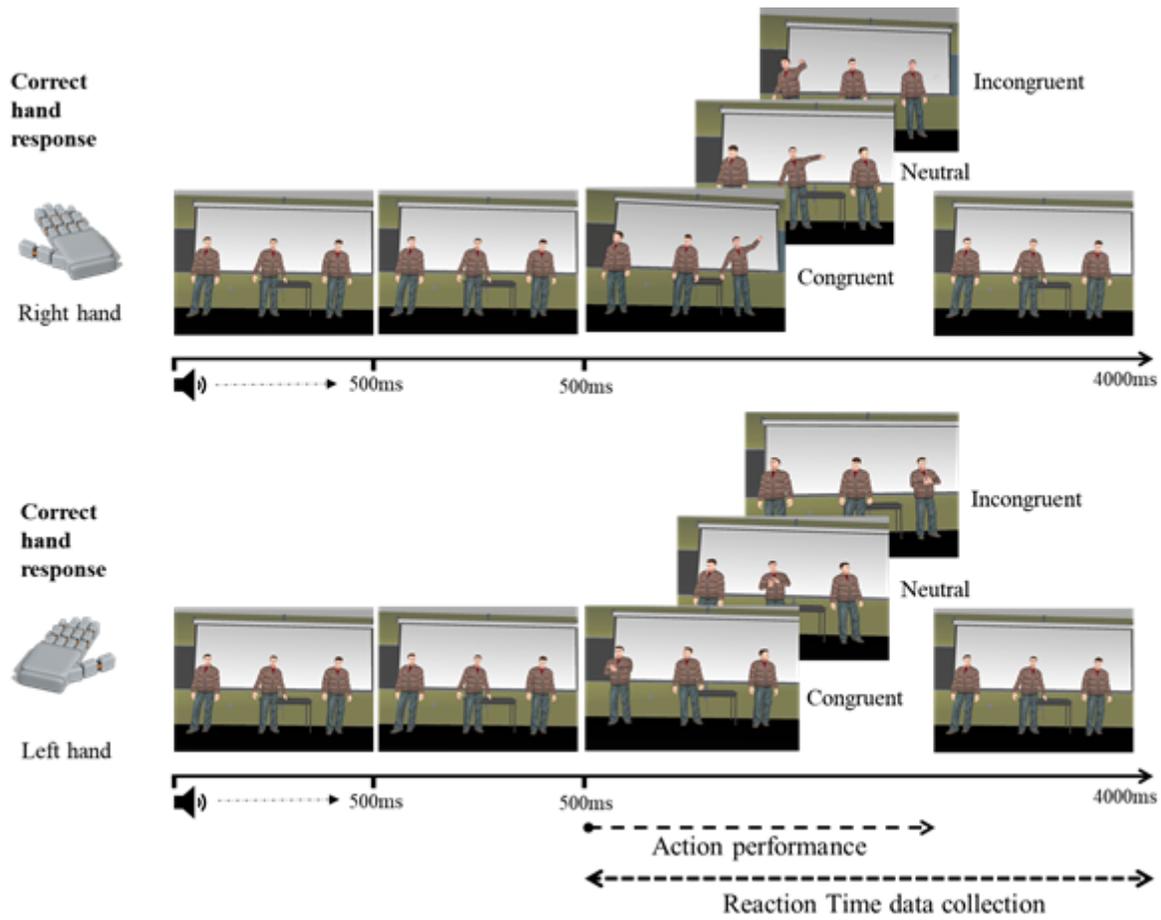


Figure 2

Time course of the 3D stimulus presentation in the VR adaptation of the Simon task.

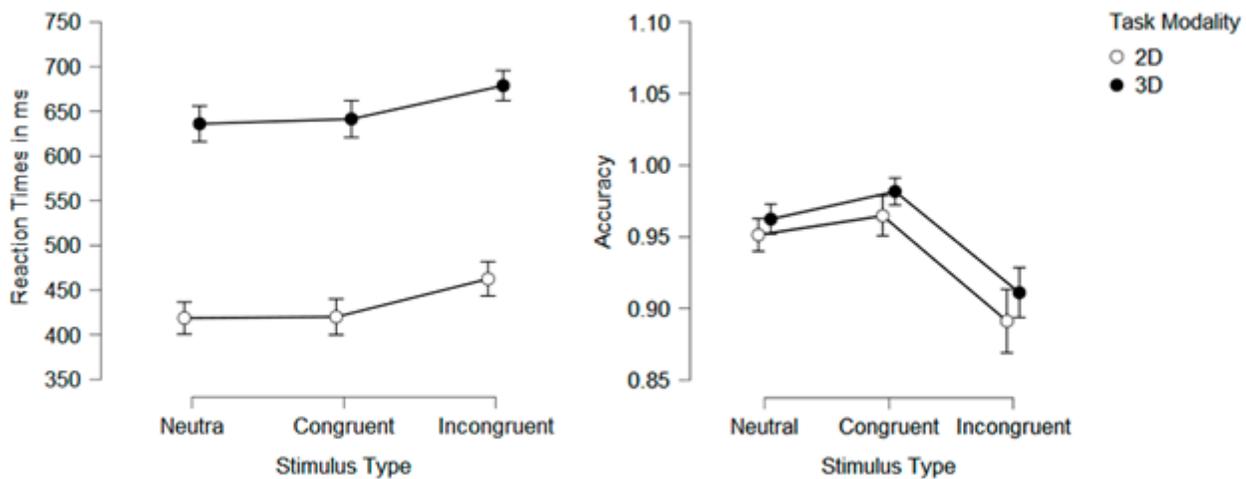


Figure 3

Mean RT (left) and accuracy (right) per Stimulus Type and Task Modality presentation conditions. Error bars represent 95% confidence intervals.

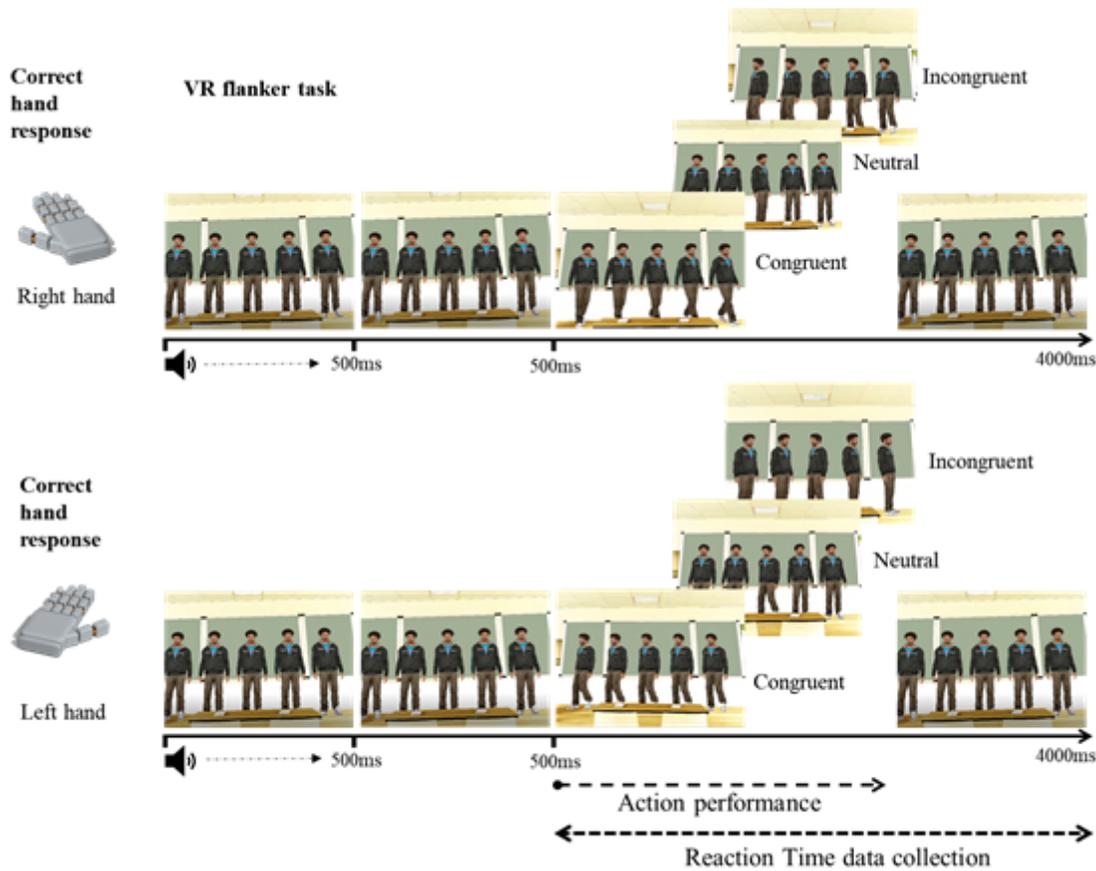


Figure 4

Time course of the 3D stimulus presentation in the VR adaptation of the flanker task.

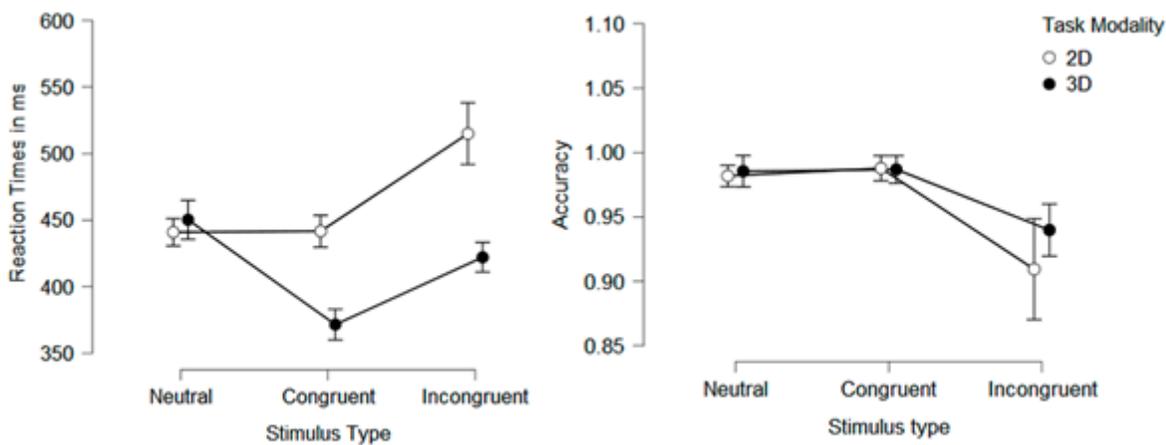


Figure 5

Mean RT (left) and accuracy (right) per Stimulus Type and Task Modality presentation conditions. Error bars represent 95% confidence intervals.