



# Switching to online: Testing the validity of supervised remote testing for online reinforcement learning experiments

Gibson Weydmann<sup>1</sup> · Igor Palmieri<sup>1</sup> · Reinaldo A. G. Simões<sup>1</sup> · João C. Centurion Cabral<sup>2</sup> · Joseane Eckhardt<sup>3</sup> · Patrice Tavares<sup>1</sup> · Candice Moro<sup>3</sup> · Paulina Alves<sup>1</sup> · Samara Buchmann<sup>1</sup> · Eduardo Schmidt<sup>1</sup> · Rogério Friedman<sup>3</sup> · Lisiane Bizarro<sup>1</sup>

Accepted: 19 September 2022  
© The Psychonomic Society, Inc. 2022

## Abstract

Online experiments are an alternative for researchers interested in conducting behavioral research outside the laboratory. However, an online assessment might become a challenge when long and complex experiments need to be conducted in a specific order or with supervision from a researcher. The aim of this study was to test the computational validity and the feasibility of a remote and synchronous reinforcement learning (RL) experiment conducted during the social-distancing measures imposed by the pandemic. An additional feature of this study was to describe how a behavioral experiment originally created to be conducted in-person was transformed into an online supervised remote experiment. Open-source software was used to collect data, conduct statistical analysis, and do computational modeling. Python codes were created to replicate computational models that simulate the effect of working memory (WM) load over RL performance. Our behavioral results indicated that we were able to replicate remotely and with a modified behavioral task the effects of working memory (WM) load over RL performance observed in previous studies with in-person assessments. Our computational analyses using Python code also captured the effects of WM load over RL as expected, which suggests that the algorithms and optimization methods were reliable in their ability to reproduce behavior. The behavioral and computational validation shown in this study and the detailed description of the supervised remote testing may be useful for researchers interested in conducting long and complex experiments online.

**Keywords** Online remote experiment · Reinforcement learning · Computational modeling

Online experiments are an alternative for researchers interested in conducting behavioral research outside the laboratory with a great number of subjects. There was an increase in the use of remote experiments for behavioral experimentation

and neuropsychological assessment in the last years. Empirical studies suggest similar results for online and in-person behavioral assessments (Brearly et al., 2017; Carr et al., 2020; Chaytor et al., 2021; Wadsworth et al., 2016), but recommendations for bona fide online experiments need to be followed in order to ensure data quality (Grootswagers, 2020; Sauter et al., 2020). During the coronavirus disease (COVID-19) pandemic, for instance, the option for remote experiments was not just a need but it was one of the only ways to conduct research without risking the lives of human subjects. With this in mind, behavioral researchers had to formulate creative solutions for online data collection during the pandemic (Bilder et al., 2020; Gagné & Franzen, 2021; Geddes et al., 2020; Tailby et al., 2020).

Practical issues in the applicability of online behavioral research were documented in the last years. As extensively noted in previous publications, researchers still needed to

✉ Gibson Weydmann  
gibson.weydmann@ufrgs.br; gibsonjw6@gmail.com

<sup>1</sup> Programa de Pós-Graduação em Psicologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

<sup>2</sup> Instituto de Ciências Humanas e da Informação, Universidade Federal do Rio Grande (FURG), Rio Grande, Brazil

<sup>3</sup> Programa de Pós-Graduação em Ciências Médicas: Endocrinologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

solve problems with the Internet connection and timing issues, prove the reliability of the measures applied, replicate online phenomena that were observed in laboratory settings, control for the occurrence of distractors during the experiment, demonstrate the usability of the employed software, and take extra care to provide clear instructions for participants (Cernich et al., 2007; Gagné & Franzen, 2021; Geddes et al., 2020; Grootswagers, 2020; Holmlund et al., 2019; Sauter et al., 2020; Tailby et al., 2020). Despite these limitations, data from studies comparing face-to-face and online behavioral data suggest that similar results are observed for both data collection methods (Brearly et al., 2017; Carr et al., 2020; Chaytor et al., 2021; Wadsworth et al., 2016).

One of the main disadvantages of online experiments is the limited experimental control that exists when subjects self-apply the behavioral measure within their natural environment (Gagné & Franzen, 2021; Grootswagers, 2020). This can be a major problem when online behavioral experiments are complex and involve many steps for data collection. Sauter et al. (2020) mention that online assessments are supposed to be short given that many additional instructions and careful steps would be needed to ensure data quality in longer (~a 1 h) experiments. Therefore, creative solutions were brought by researchers to ensure data quality in longer and complex experiments. Wadsworth et al. (2016), for instance, created a videoconference neuropsychological assessment that mimics the usual face-to-face procedures in order to conduct many assessments to a culturally diverse sample of rural American Indians. The authors compared remote (i.e., videoconference) and face-to-face assessments and found that both methods had similar results for the population tested. More recently, Cuttler et al. (2021) used a remote assessment method to apply a series of neuropsychological tests in a highly controlled experiment on the effect of cannabis use. The authors used a Zoom® link to control the use of cannabis during the experiment and to ensure the comprehension of task instructions and engagement in neuropsychological testing, all essential for the execution of the experiment. Finally, the results from the meta-analysis conducted by Brearly et al. (2017) suggest that synchronous online neuropsychological assessment might be needed when experiments are complex. The authors found reliable outcomes when online assessments were conducted in real time with the remote presence of the experimenter/health care worker and without interruptions in communication.

Although the use of synchronous assessments seems counter intuitive as it reduces the freedom of online assessment, its use might be a solution for experimenters interested in running complex and long online experiments with multiple steps (e.g., the experiment from Cuttler et al., 2021). Subjects are also not allowed to repeat tasks and habituation and attention can be controlled in synchronous experiments. Furthermore, research outside the field of neuropsychology

indicates that people may be more prone to fake online psychological assessments when they are not supervised because the acceptance to do so (i.e., the subjective norm) online is higher than doing it in person (Grieve & Elliot, 2013).

One sure way to assume that the online experiment was conducted correctly is to test the association between online and in-person data. This has been done extensively in neuropsychological research using normative tasks well validated in the literature (e.g., n-back for working memory), but behavioral measures that assess differences in learning mechanisms have rarely been considered in those studies. For instance, apart from the increasing number of studies about reinforcement learning (RL) (Adams et al., 2016; Huys et al., 2021), online experiments with reinforcement tasks seem to be scarce. One exception is the research by Nussenbaum et al. (2020), in which the authors were able to replicate online the effects of age on RL observed in two in-person studies.

In the last two decades, there has been an increase in the use of RL tasks to assess abnormal reward and punishment learning in people with psychiatric disorders (Adams et al., 2016; Collins et al., 2014; Huys et al., 2021). RL has been studied in psychology since the early works of Edward Lee Thorndike and Burrhus Frederic Skinner and refers to the ability to maintain or change behavior due to its consequences in the natural environment (Donahue, 2017; Sutton & Barto, 2018). Positive reinforcement of a behavior is the process by which a behavior is maintained because it generates rewards/positive feedback (i.e., win-stay responses), whereas negative reinforcement refers to the ability to change behavior to remove aversive states or avoid negative feedback (i.e., lose-shift). Case-control studies using RL tasks have revealed promising and interesting results regarding individual differences in RL, such as the finding that impairments in learning by rewarding consequences on schizophrenia might be more related to lower working memory (WM) utilization in trial-and-error learning than to problems with learning by feedback (Collins et al., 2014). Recently, RL has even been also studied developmentally, with data showing that the ability to learn by reinforcement improves from an early age (8 years old) to adulthood (Master et al., 2020; Nussenbaum et al., 2020).

A big advantage of RL studies is the use of computational models to simulate brain activity using trial-by-trial data. Computational models of psychological phenomena are formal causal explanations of brain function that occupy a central position in the generative causal models of computational neuroscience (Huys et al., 2011; Huys et al., 2021). RL has been studied with computational models since the 1970s and many algorithms have been created since then to replicate laboratory findings using machine learning environments (Sutton & Barto, 2018). Because computational models are expected to replicate the computations that the brain does when solving real-world problems, algorithms

designed to explain behavioral phenomena need to predict real behavioral data to be valid. Furthermore, when more than one computational model exists for the same data, they need to be statistically compared to find out which one is the best to predict behavioral phenomena. These assertions were the basis for the proposal of computational validity, that is, the understanding that computational models can be valid when they predict the behavior of different species or distinct experimental manipulations (Redish et al., 2022). Therefore, a model for reinforcement learning can be valid when it predicts behavior under distinct conditions, for instance, as is the case of online and in-person experiments.

Based on the aforementioned, this paper aims to test the computational validity of a supervised remote testing RL experiment conducted during the pandemic. The option for a remote and synchronous experiment was done because the original experiment was planned to do in person and involve multiple steps to be followed in a specific order. Based on the idea of computational validity (Redish et al., 2022), it was assumed that our remote experiment and the behavioral and computational results would be valid only if they would replicate previous data using the same behavioral task. The feasibility of our remote and synchronous online experiment and the usefulness of low-cost and open-source programming languages to analyze data are discussed in this paper.

## Method

### Participants and study design

This study is part of a bigger project titled Health Young at Risk for Obesity (HYRO) approved by the Ethics Committee of Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Brazil (register number: 62798716.2.0000.5327). The aim of the HYRO project was to investigate how risk factors for obesity impact the eating behavior of normal-weight (body mass index = 18.5 to 25 kg/m<sup>2</sup>) young adults (18 to 24 years old). The project had three phases: 1) an online initial screening designed to assess risk factors for obesity, socio-demographic characteristics and dieting and psychological factors; 2) an in-person assessment at the Institute of Psychology at Universidade Federal do Rio Grande do Sul (UFRGS) designed to evaluate inhibitory control and RL using behavioral experiments, to collect anthropometric and to assess dieting profile using a food-frequency questionnaire; 3) a blood sample extraction at HCPA for posterior metabolic analyses. The exclusion criteria for the original project were being outside the age or BMI range.

Phase 1 of the project HYRO was conducted between 2020 and early 2021 and the final sample for this phase was 420 young adults eligible for the next phases. All participants gave informed consent for the online assessment and

were informed that they could be called to be at the Institute of Psychology at UFRGS for a series of behavioral assessments or at HCPA for blood extraction. However, due to the COVID-19 pandemic and social-distancing restrictions implemented in March 2020 in Brazil, procedures for the in-person assessments were adapted to an online format. Phase 2 was then divided into two online phases that were conducted concomitantly: a remote and synchronous behavioral experiment designed to assess RL and impulsivity, and an online, guided self-report assessment of eating habits.

For the conduction of the remote behavioral experiment, an additional exclusion criterion was then applied: the participants had to have access to a laptop or a desktop computer to be able to perform the assessments. All 420 subjects from phase 1 were contacted for phase 2. The majority of participants called to this phase accepted to participate ( $n = 242$ , 57.70% of the total sample of  $n = 420$ ); 177 (42.14%) participants did not answer the invitation, ten (2.38%) participants declined the invitation, three (0.71%) participants were unable to find a date/hour for the experiment, and one (0.23%) participant had the experiment interrupted and his data deleted because his responses were considered slower and random by the experimenter.

For the current study, 30 subjects that participated in our remote experiment (nine men; age  $M = 20.83$  years,  $SD = 1.62$ ) were randomly selected from the bigger sample of 242 participants using the `random.sample` function in Python programming language (Python foundation, <https://www.python.org/>). We chose to randomly select participants to avoid selection biases since our sample included subjects with distinct risk factors for obesity and with diverse psychological profiles. Also, the relationship between these individual differences and RL performance are being investigated using a hypothesis-driven approach unrelated to the aim of the present study. The number of subjects selected represents about 10% of our total sample.

**Reinforcement learning/working memory task (RLWM)** This task was adapted from Collins (2018) to assess how WM load impacts RL performance. The task had two phases: a RL phase in which participants had to learn which button to press for visual stimuli, and a memory phase in which participants needed to recall their responses for the stimuli of the first phase. In the present study, only the data for the RL procedure will be presented because of its frequent replication in the literature (Collins, 2018; Collins & Frank, 2012; Collins et al., 2014; Master et al., 2020).

Participants received an instruction to learn, by trial and error, which of three buttons (*c*, *b*, and *m* on the computer keyboard) they needed to press for each stimulus they saw on their computer screen. They were explicitly instructed to use correct and incorrect feedback as a guide to keep or change their response to the visual stimuli. Feedback was

presented for each response following a continuous schedule of reinforcement. The task had 14 blocks and WM load was controlled by the number of stimuli presented on each block. In the low WM load condition (eight blocks), three stimuli were randomly presented to each participant, while in the high WM load condition (six blocks) six stimuli were presented in a random order (Fig. 1). For each block, the discriminative feature of stimuli (e.g., color, geometric figures, cartoons, landscapes, etc.) was changed to ensure that participants were engaging in a new learning process. Previous studies showed that the number of stimuli presented in each block influences the acquisition of learning due to the number of stimulus–response–feedback associations needed in each condition: on low WM load condition, RL is influenced by WM performance because participants might remember their response in the outcome for  $n$  trials before, while in the high load condition RL is not influenced by WM due to overload of its capacity (Collins, 2018).

The first and last blocks of the task were of low-WM load. Each stimulus was presented nine times in a total of 540 trials. Each trial was composed of a cross-fixation that lasted for 500 ms followed by a visual stimulus presented until participants responded and then by the feedback screen for 1000 ms. Participants had 24 training trials using six stimuli unrelated to the main task before the real

assessment. As done in previous studies with the RLWM, a performance criterion was established for their data to be analyzed: Participants had to show an average accuracy higher than 75% on the last two presentations of each stimulus. All participants randomly selected for this study had an overall performance above this criterion ( $M = 82%$ ,  $SD = 4.9%$ , range, 78–98%).

**Procedure** The behavioral experiment was initially planned to run in the laboratory. Research assistants would monitor participants throughout the process and schedule a date and time for each participant, with a mean session time of 1 h. The plan for the original experiment was: a) subjects would start the initial phase of the RLWM task; b) complete an inhibitory control task; c) do the final phase of the RLWM before starting anthropometric and dieting assessments. On switching to online, the essence of the original procedures of in-person assessment was maintained. Therefore, the option of a remote, synchronous online contact with participants was selected in order to mimic the original idea; research assistants were available online to monitor the entire data collection. The monitoring was useful in this case because our in-person and online pilot studies revealed that some participants were more prone to disengage from the task due to its longer duration (~1 h).

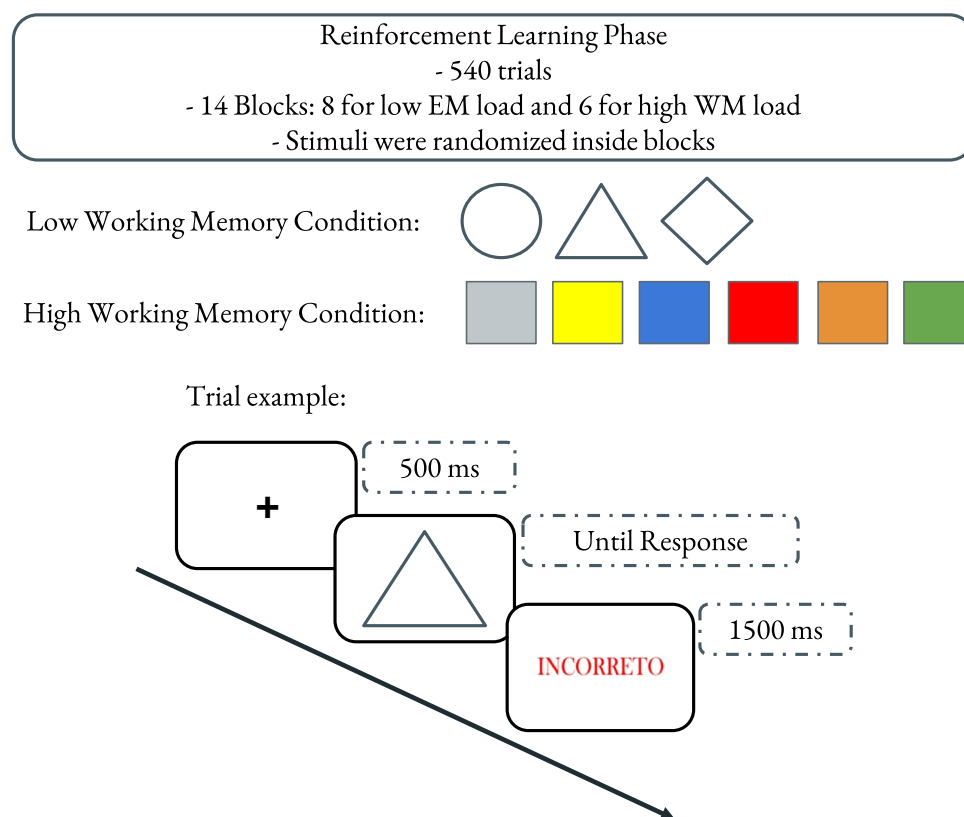


Fig. 1 Reinforcement Learning Working Memory (RLWM) task structure

All participants invited to this study were contacted via mobile phone and they agreed on a convenient date and time between 8 AM and 8 PM. Participants were advised to stay at home during the experiment and received instructions to be alone in a room without distractors, and to remain focused on the task. They provided a new informed consent for the remote experiment.

Each experimenter had his/her own computer for data collection, and each computer had the psychological experiments software OpenSesame® (version 3.3.2 for Windows 64 bits; Mathôt et al., 2012) installed. Before starting the contact with participants, each experimenter had to train for the remote experiment and follow the established procedures without error with at least two volunteer subjects unrelated to the project. Standard operating procedures to report and solve problems in internet connection and task performance were provided by the first author.

On the scheduled day and time, participants received instructions through mobile app messages (via WhatsApp® or Telegram®). We noticed in a pilot study that sending the message via mobile was more engaging for participants because they use WhatsApp/Telegram to communicate in their day-to-day. More importantly, changing messages with the experimenter during the assessment was a way to keep participants active in the experiment since each step in the remote testing was started by them. In order to guarantee the execution of the procedures, instructions were divided into five steps (A to E) and experimenters were online and available to answer questions and guide subjects all the time. A summary of the procedures is shown in Fig. 2.

Experimenters were trained to ensure that participants were following the instructions for each step before continuing the data assessment, and the listed interferences in Fig. 2 (dashed lines) would cause the experiment to end. Instructions about Google Remote Access (GRA®, version 1.5) were provided for the remote access. GRA is a free tool used to access geographically distant computers from your own computer using a network connection. In our case, participants were invited to access the experimenters' laptops using a GRA link. Then they received a password provided by the experimenter and used it on the site to access the experimenter's computer. With remote access from GRA, participants were able to perform from their own home the behavioral tasks that were running in the experimenters' laptop without installing any software on their own machine. Importantly, the experimenter did not have access to the participant's computer.

After starting remote access, participants were able to control the RLWM through their own computers using their Internet browser. The remote access allowed the experimenter to see everything that participants saw on their screen (i.e., instructions, stimuli, and feedback). The option to monitor subjects' performance and see what they see along the task is useful

because the researcher can monitor the participants' performance and identify problems with the Internet connection or careless responses. As it happens on synchronous neuropsychological assessments, experimenters were instructed to monitor participants' performance and experiment time and register potential outliers on the lab book. The experimenters were able to identify one subject who had problems with the Internet connection during the experiment, two participants who forgot to resume the experiment after the first phase of RLWM, and one participant who took a much too-long interval between tasks (more than 15 min). These four participants did the first part of RLWM before the issues were mentioned, and were still eligible to be selected for the present study. After the first phase of RLWM was finished, remote access was stopped and started again for another behavioral task and then for the second phase of RLWM.

## Data analysis

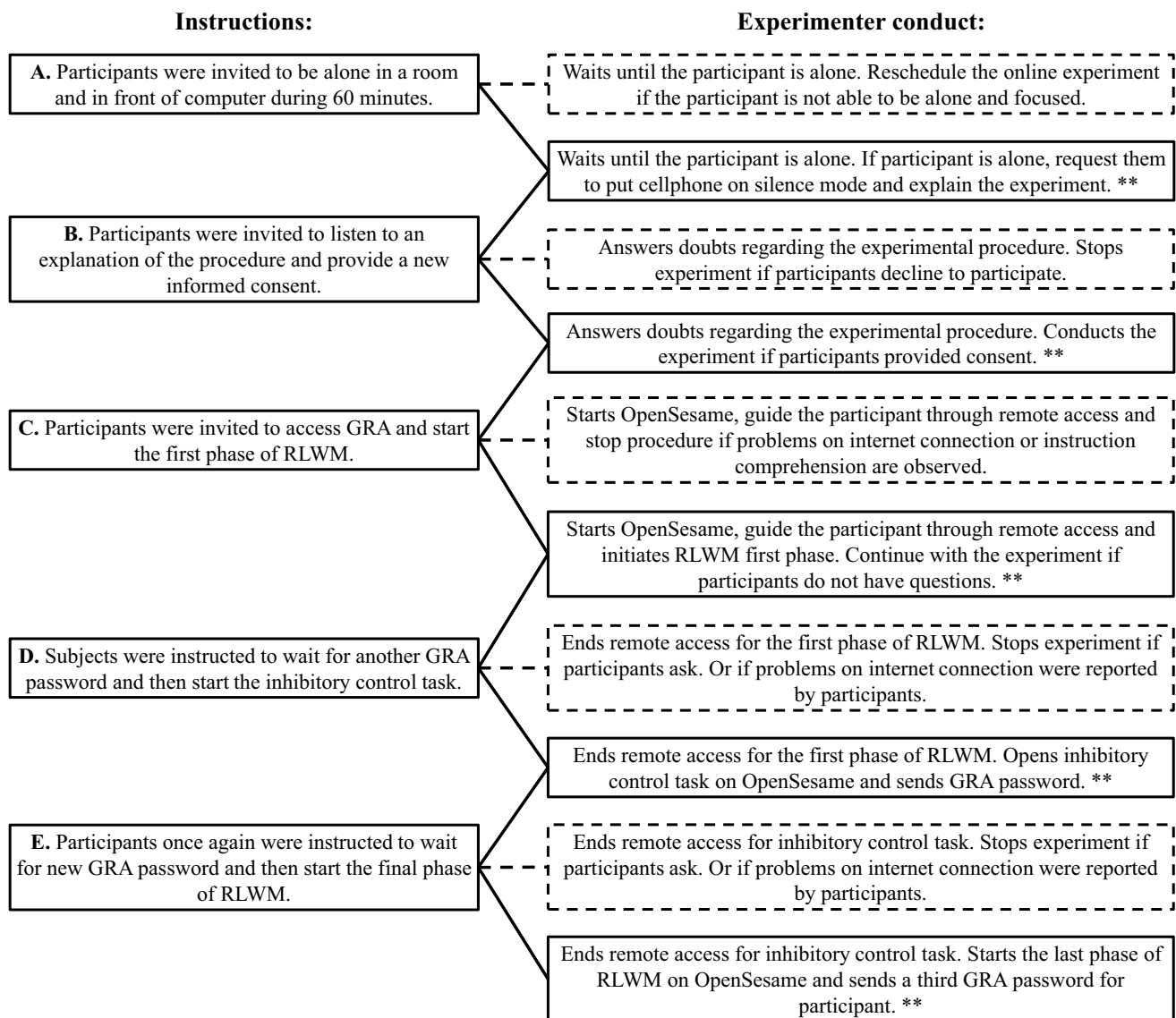
**Programming tools and statistical testing** All the codes responsible for data processing, visualization, and computational modeling were implemented using a Python 3.9 environment. The most relevant libraries used in this project were pandas 1.2.3 (McKinney et al., 2010), for data manipulation and organization, and scipy 1.6.1 (Virtanen et al., 2020), for parameter optimization and auxiliary model functions. This code is organized in a git repository, which contains a complete list of its dependencies. We also made extensive use of the free Google Colaboratory® infrastructure to run code for data analysis, model training and simulations.

Following previous publications with RL tasks, we divided our analysis into model-free and computational modeling (Collins, 2018; Daw, 2011). The model-free analysis involves using statistical tests to evaluate to what extent task-related procedures impact behavioral performance.

All data were analyzed using trial-by-trial records in order to assess how WM load (set size 3 or 6), stimulus iteration, delay, positive reinforcement history (PR), and blocks number impact chances of correct responses and RT. Multiple regressions were conducted then for correct/incorrect feedback (binary logistic model) and for RT (linear model). Because our supervised remote testing relied on an Internet connection, we assumed that problems with RT latency and distribution could have happened. RT distribution can be higher for online experiments (e.g., Semmelmann & Weigelt, 2017) and pre-buffering the entire experiment or relying on online software for psychological experiments might reduce this noise (Grootswagers, 2020; Sauters et al., 2020). In our case, however, these options were not a choice since online software for psychological experiments are recommended for short assessments, which was not the case in our study (Sauter et al., 2020).

On operant learning procedures, a decrease in RT is expected as a function of learning because individuals are





**Fig. 2** Instruction and experimenter conduct for the remote behavioral experiment. The instructions were divided into five steps (A to E) and each step needed to be finished before going to the next. *Dashed lines* indicate cases in which experimenters would need to stop data

collection. \*\* Indicates that experimenters were registering the process in the lab book. *GRA* Google Remote Access. *RLWM* Reinforcement Learning and Working Memory task

becoming faster in their response after the response-outcome associations. Therefore, an inspection of our data should reveal the expected effects of learning over RT (see Figure 2 on Collins, 2018 for an example). To verify whether RT was decreasing as a function of stimulus presentation, a graph for the interaction between mean group RT and stimulus iteration was plotted (Supplementary Material - Figure S1A) and revealed that participants had the expected performance through the task even when the raw RT was considered: A decrease in RT was observed as stimulus iteration increased, and high-WM load blocks had higher RT than low-WM blocks. However, RT distribution was not normal. To remove response outliers, we deleted RT slower than 200 ms or higher than the

mean plus three standard deviations for each subject (a total of 1.99% [323 out of 16200] responses were deleted and a maximum of 3.33% responses were deleted for a single subject). Supplementary figure S1B shows that corrected RT still had a non-parametric distribution (Figure S1C). Log transformation was then applied and this transformed score was included as an outcome on linear regression because of its parametric distribution (Supplementary Figure S1D).

The predictors chosen are supposed to reflect the effects of WM and learning through the task. WM load represents the expected effect of set size over RL. Delay also indicates how WM impacts task performance and it is calculated using the number of trials between a correct response for stimulus

$x$  and the next time a correct response for  $x$  occurs. The variable PR counts the number of previous correct choices for a given stimulus, and the block number counts the number of blocks from the beginning to the end of the task. While WM load and delay are expected to negatively impact the chances of correct response and increase RT, PR, and block number are expected to increase correct responses and decrease RT. All predictors were transformed into  $z$  scores before regression analysis and influential cases were controlled using Cook's distance  $4/n$  criteria. All statistics were conducted using R programming language and RStudio (R Core Team, 2020; RStudio Team, 2019).

**Computational modelling** Computational modelling was used to fit subject behavior data and to study the effects of RL and WM in task performance. Four candidate models were tested, all based on previously published models and algorithms (Collins, 2018; Collins & Frank, 2012; Collins et al., 2014; Master et al., 2020) to verify the computational validity.

**Classic RL** The original models were built upon a simple RL algorithm. The classic RL is the simplest model as described by Sutton and Barto (2018):  $Q_{t+1}(s, a) = Q_t(s, a) + \alpha_{RL} \delta_t$ . Including the two-parameter Q-learner, which updates each learned value, that is, the expected reward “ $Q$ ” for the selected action “ $a$ ”, given the stimulus “ $s$ ” upon each trial's reward outcome “ $r_t$ ”, where  $r_t = 1$  for correct or  $r_t = 0$  for incorrect. Choices were generated probabilistically and actions were selected with likelihood according to higher Q-values. The rule for choosing actions in response to a stimulus was defined stochastically by a Softmax Choice Policy:  $P(as) = \exp(\beta Q(s, a)) / \sum_i \exp(\beta Q(s, a))$ , where  $\beta$  is the inverse temperature parameter which determines the degree to which differences in Q-values are translated into a more deterministic choice; the sum is over the  $n_A = 3$  possible actions and Q-values were initialized to the uniform random policy  $U = 1/n_A$ . The  $\beta$  parameter varies significantly from one study to another (e.g., 0 to 100 or 0 to 500) and some authors suggest constraining its range as a free parameter or adopting small values in order to achieve a good model fit (Daw, 2011; Wilson & Collins, 2019). Based on previous studies using the RLWM task (Masters et al., 2020), we decided to fix  $\beta$  as 50 because parameters distribution was close to normal in this way and because free beta ranges (e.g., 0 to 50 or 0 to 500) lead to extreme values of beta in many cases with minimal or no change on function cost.

**Additional parameters ( $\epsilon$ ,  $init$ ,  $pers$ ,  $\varphi_{RL}$ )** Classic RL assumes that participants will learn from trial and error, but biases might influence participants performance and additional computational parameters might control response bias. These parameters are included in the models RLWMI and pure WM, shown in the next section.

An example of response bias is that participants may prefer to follow a certain order on their choices or choose to always press a specific button at the beginning of each trial. To control for this type of response an “initial bias” parameter titled *init* was created. In this parameter, the first action selected by the participant for each image they see acts as a marker of a potential bias and then boosts the value of this choice before the first learning rate update. The following formula is followed for the *init* parameter:  $Q_0(s, a_{first}(s)) = 1/n_A + init * (1 - 1/n_A)$ .

Slips of action are also possible and we captured in an undirected noise parameter,  $\epsilon$ , which is a free parameter independent from learning that represents the amount of noise in the data: the agent chooses an action based on the softmax probability with the probability  $1 - \epsilon$ , and lapses or chooses randomly with the probability  $\epsilon$ . The definition of the new mixture choice policy becomes:  $P(s|a)' = (1 - \epsilon) P(s|a) + \epsilon U$ .

Forgetting the previous responses may also influence performance. The potential decay at each trial toward the initial uninformed  $Q$ -value was modeled following the formula  $Q_0 = 1/n_A$ :  $Q_{t+1} = Q_t + \varphi_{RL}(Q_0 - Q_t)$ , where  $0 < \varphi_{RL} < 1$  is the forgetting parameter for the RL model, named decay.

Finally, on the RLWM task, some participants might be neglecting negative feedback and show perseverative errors despite negative outcomes. The persistence parameter (*pers*) was modeled to control perseverative responses and it represents a positive learning bias parameter so that the learning rate  $\alpha$  is maintained in positive prediction errors ( $\delta \geq 0$ ), and it is reduced by  $\alpha_{RL} = (1 - pers)$ . Values of *pers* that equal 0 indicate equal learning from positive or negative feedback, whereas values that equal 1 indicate complete neglect of negative feedback, that is, there is a tendency to repeat a behavior that led to negative feedback.

**WM model** The pure working memory model (Collins & Frank, 2012; Collins et al., 2014) performs rapid updating but capacity-limited WM to learn stimulus action values  $W(s, a)$ . In this model, participants' performance is supposed to be influenced only by WM. Fast learning is represented by  $\alpha_{WM} = 1$ , and outcome value is  $r_t = 1$  for correct or  $r_t = 0$  for incorrect responses. This model includes a probabilistic capacity limitation to attenuate the working memory effect for blocks containing more than  $K$  stimuli. In our case, we supposed that the best scenario would be to assume that participants were able to remember all six stimuli from the high-WM load condition, and thus  $K = 6$ . For the WM model, the effect of memory over performance is represented by the variable  $\eta_{WM} = \eta \cdot \min(1, K/n_s)$ , where  $\eta$  is a policy mixture parameter. The model also considers the influence of the additional decay parameter ( $\varphi_{WM}$ ), and shares with the RL model the parameters persistence (*pers*), noise ( $\epsilon$ ) and the inverse temperature ( $\beta$ ). The overall choice policy that expresses the WM-only

involvement in choice becomes  $P(als) = \eta_{WM}P_{WM}(als) + (1 - \eta_{WM})P_{Other}(als)$ , where  $P_{WM}(als)$  is the softmax adjusted policy over  $W(s, a)$  including all mechanisms described above, and  $P_{Other}(als) = U$ .

**Reinforcement learning and working memory interaction model (RLWMI)** The RLWMI model was created using as a reference the RLWM model. The RLWM model assumes that information stored for each stimulus in WM also pertains to action–outcome associations and includes two independent and non-interacting mechanisms of learning at the level of choice. The first mechanism is an RL module similar to the classic RL, including as parameters the learning rate  $\alpha_{RL}$  and the softmax inverse temperature  $\beta$ . However, the RL module model also considers the additional parameters  $\epsilon$ , *init*, *pers*, and decay ( $\varphi_{RL}$ ).

The second mechanism of the RLWM is a WM module that includes fast storage of information according to weights between stimuli and actions using the formula  $W_{t+1}(s_{t+1}, a_{t+1}) = r_t$ . The WM module is initialized similarly to the previous RL Q-values but captures perfect retention of the information from the previous trial so that the learning rate is  $\alpha_{WM} = 1$ . As WM maintenance tends to fail increasingly in time and with the intervention of other stimuli, we assumed that WM is delay-sensitive and so weights also decay on each trial according to the formula  $W_{t+1} = W_t + \varphi_{WM}(W_0 - W_t)$ . The WM module shares three parameters with the RL module of the RLWM model (i.e., softmax choice  $\beta$ , the indirect noise parameter  $\epsilon$ , and response bias parameter *pers*). Finally, the limitations of WM use on the task were modeled using a two-mixture parameter that considers the WM involvement on the low- ( $\eta_3$  for set size 3) and high- ( $\eta_6$  for set size 6) WM load conditions through the formula  $P(als) = \eta_{ns} P_{wm}(als) + (1 - \eta_{ns}) P_{RL}(als)$ .

The RLWMI is the same as the RLWM model, with the exception that the WM module influences the RL computations (Collins, 2018). The RL module still follows the Q-learning equation from the classic RL model, however, WM contributes to the computation of the reward prediction error  $\delta_t$  according to the proportion to WM's involvement in choice, where  $\delta_t = r_t - (\eta_{ns} W_t(s, a) + (1 - \eta_{ns}) Q_t(s, a))$ . The main contribution of the RLWMI is the prediction on the cooperation of the interaction between both modules, including opposite effects of set size in learning, in which performance is worse in higher set sizes.

**Model optimization** Once implemented, models must be instantiated with numeric values for parameters to be able to reproduce an individual behavior during the experiment. For each model and individual pair, there is a set of values that makes the model output as close as possible to the data observed, and the task of finding these values is considered a model parameter optimization problem.

The first step in this task is to define a cost function that would enable a quantitative measure of the fitness of each model equipped with a given set of parameter values. Considering the maximum likelihood criterion, one could start with a probability of the experiment answers, which for an individual whose output is given in  $d_{1:T}$ , and a model with a set parameters, can be directly derived from its policy:

$$L(\theta|d_{1:T}) = \prod_{t=1}^T p(c_t|d_{1:t-1}, \theta),$$

that is, the probability of this model taking the choices  $a=c_t$  at each respective trial. Introducing the logarithm to avoid numerical stability problems, and to work as a cost minimization task, the final function to be optimized can be defined as

$$J(d_{1:T}, \theta) = -\ln[L(\theta|d_{1:T})] = \sum_{t=1}^T \ln[p(c_t|d_{1:t-1}, \theta)].$$

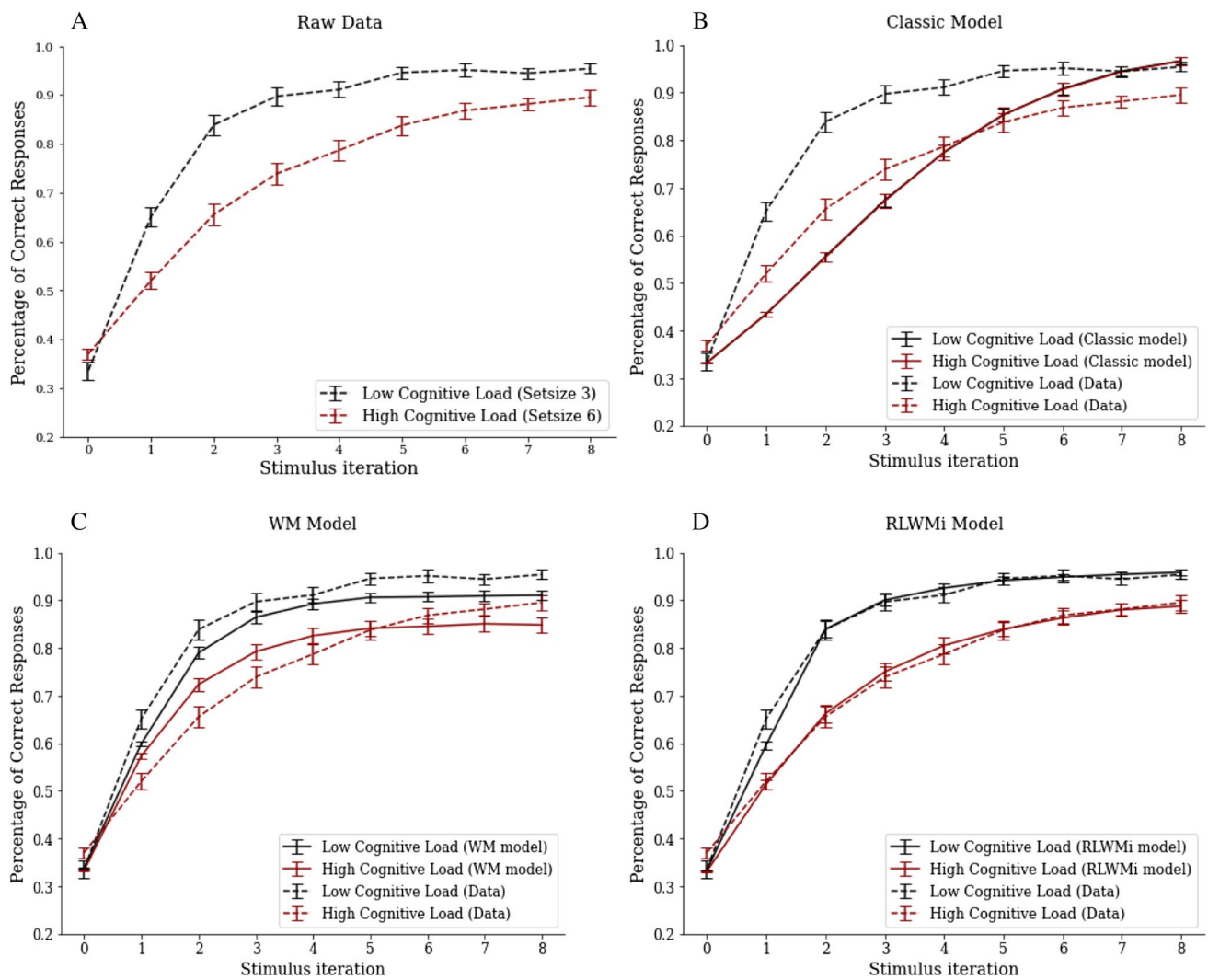
This function can be easily calculated with the model policy and data, and was then minimized using the *scipy.optimize.minimize* tool, configured for bounded optimization of its free parameters. The optimization procedure was repeated 20 times for each case, with random initial values, to avoid local minima.

## Results

### Model free

Learning curves were plotted considering the effects of WM load condition over learning. Figure 3A reveals that participants had the expected learning curves. As iteration increases, the mean percentage of correct responses increases as a function of WM load: under low cognitive load, participants had improved learning, while under high-load condition, an attenuation in the learning curve was revealed. Multiple logistic regression analysis confirmed the effects of WM variables and learning predictors. WM load had a negative effect over learning ( $B = -0.421$ ,  $SE = 0.047$ ,  $p < 0.001$ ), with the chances of making a correct response decreasing on set size 6. Delay also decreased the chances of correct responses, meaning that the number of trials between correct responses for the same stimulus impaired learning ( $B = -0.508$ ,  $SE = 0.028$ ,  $p < 0.001$ ). PR had a positive effect on learning, with the chances of making a correct response increasing as participants made correct choices for the same stimulus ( $B = 2.262$ ,  $SE = 0.043$ ,  $p < 0.001$ ). Block number also increased the chances of correct responses ( $B = 0.136$ ,  $SE = 0.035$ ,  $p < 0.001$ ), meaning that





**Fig. 3** Learning curves as a function of stimulus iteration for each condition and correspondence between computational models and raw data. Simulations for the computational models were executed 100 times for each subject (see Model optimization for more details).

**A** Learning curves for each condition. **B** Correspondence between Classic RL model simulation and raw data. **C** Correspondence between a pure WM model simulation and raw data. **D** Correspondence between RLWMI model simulation and raw data

performance was improving throughout the task. Importantly, predictors had an opposite effect on log-RT, also confirming the expected results. WM load condition ( $B = 0.139$ ,  $SE = 0.005$ ,  $p < 0.001$ ) and delay ( $B = 0.046$ ,  $SE = 0.004$ ,  $p < 0.001$ ) increased RT in the task, showing that WM load might influence slower reactions. The two other predictors were associated with performance improvement because both predicted faster RT in the task: PR ( $B = -0.139$ ,  $SE = 0.004$ ,  $p < 0.001$ ) and block number ( $B = -0.032$ ,  $SE = 0.004$ ,  $p < 0.001$ ).

### Computational modeling

Of the three computational models tested, RLWMI had the best fit according to AIC (see Table 1). This result was also

observed when learning curves using model simulations were created, as shown in Fig. 3. The classic RL model (Fig. 3B) was unable to capture learning differences related to cognitive load, as can be seen by the overlap between learning curves. While the pure WM model captured differences in performance that are related to the WM conditions, this model was still insufficient to account for the ascendant proportion of correct trials in the last iterations (Fig. 3C). Finally, learning curves obtained with the RLWMI model were closer to the ones obtained with real subjects, representing both the effects of WM condition and the acquisition of learning from the first to the last stimulus iteration. To further validate the RLWMI model, a logistic regression using values obtained with 100 simulations for each subject

**Table 1** Mean and standard deviation for model parameters

Models	Parameters							
	<i>Learning rate</i> ( $\alpha$ )	Decay ( $\phi$ )	Persistence ( <i>pers</i> )	Random noise ( $\epsilon$ )	Initial bias ( <i>init</i> )	$\eta^3$	$\eta^6$	AIC
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Classic	.03 (.008)	—	—	—	—	—	—	638.14 (128.00)
WM	—	.20 (.058)	.85 (.069)	.02 (.043)	—	—	.90 (.074)	608.81 (137.46)
RLWMI	.11 (.168)	.07 (.034)	.68 (.256)	.04 (.039)	.01 (.011)	.59 (.265)	.18 (.192)	592.03 (134.64)

*M* mean. *SD* standard deviation. For all models,  $\beta = 50$ . For the WM model,  $\alpha = 1$ ,  $K = 6$ , and  $\eta^6 = \eta$ ;  $\eta$  is the estimated use of working memory on the task to obtain the expected behavioral outcome; *init* was not considered in the WM model. For the RLWMI model,  $\eta^3$  and  $\eta^6$  indicated the use of working memory on low and high working memory conditions, respectively. *AIC* Akaike information criterion, *WM* Working Memory Model, *RLWMI* Reinforcement Learning and Working Memory Interaction model.

was executed. The results of the simulations replicate the ones obtained with the real data for the chances of making a correct response ( $R^2$  Nagelkerke = .49), with significant effects of WM load ( $B = -0.790$ ,  $SE = 0.001$ ,  $p < 0.001$ ), delay ( $B = -0.293$ ,  $SE = 0.006$ ,  $p < 0.001$ ) and PR ( $B = 3.133$ ,  $SE = 0.006$ ,  $p < 0.001$ ). Table 1 provides the descriptive statistics for each parameter obtained with the computational models).

## Discussion

This paper tested the behavioral and computational validity of a remote and synchronous RL experiment using free software and an online methodology. Each step of the experiment was designed to provide a reliable and feasible remote assessment for researchers interested in conducting long and complex online behavioral experiments. The procedures applied in the present study follow the main recommendations from previous task forces for online neuropsychological research, as such: participants received online informed consent with specific instructions for the remote assessment and information regarding the security of their data; only participants with the appropriate technological hardware (e.g., work desktop or laptop) were eligible to participate in the study; recommendations were provided to be at a quiet place with no distractions; identity was confirmed using information (e.g., name, phone number, and e-mail) consented to by participants at an early phase of the study; instructions were provided on what to do in case of Internet connection problems; finally, participants were followed at each step of the experiment (see Box 2 from Tailby et al., 2020, for an example of a similar procedure).

The choice for a remote assessment with a scheduled date may seem counterintuitive, given that one of the main advantages of online experiments is self-application. However, our method was designed to ensure that participants were engaged in all steps of the experiment for a long period

of assessment (~ 1 h). A similar approach has already been taken in other studies in which neuropsychological tests were applied through videoconference (Cuttler et al., 2021; Wadsworth et al., 2016). Importantly, in the present study, participants self-applied the behavioral task, while the experimenters passively observed the responses on-screen, without any videoconferencing. This choice was made to attenuate observer effects, but the participants were aware that their response was being observed and this may have biased the data.

One way to assess the feasibility of new data collection methods is to compare their results with results from previous studies using other data collection methods. Face-to-face and online/remote behavioral assessments have previously been compared in the psychological literature, and results suggest that both are reliable (Brearly et al., 2017; Carr et al., 2020; Chaytor et al., 2021; Wadsworth et al., 2016). Since this comparison was not possible in this study, the only way to verify the feasibility of our assessment method was to test its capacity to replicate previously published behavioral phenomena. Different versions of the RLWM have been applied in many studies in recent years: in all cases, WM load impaired or improved RL performance as a function of the number of stimuli presented in the task blocks (Collins, 2018; Collins & Frank, 2012; Collins et al., 2014; Master et al., 2020). To our knowledge, all studies were applied face-to-face, and the behavioral data were always analyzed considering both model-free and computational methods. Our results indicate that we were able to replicate the effects of WM on RL observed in other studies and both the model-free and the computational analysis results observed here were in line with previous data (Collins, 2018; Collins & Frank, 2012; Collins et al., 2014; Master et al., 2020). In addition, we could also observe the expected effects of WM and RL on RT responses in the task adapted for this study (e.g., with fewer trials per block) and applied online. This suggests that performance and timing issues sometimes observed in online behavioral assessments could not disrupt the expected effects of WM load and RL. Regarding timing issues, we assumed that response

outliers were attenuated when the mean RT was considered since our reduced version of the RLWM task still had many trials (540). Nonetheless, a comparison between in-person and remote assessments with the same version of the RLWM would be useful to clarify whether the assessment method significantly influenced RT responses and would also contribute to the assumption of computational validity.

Finally, another potential contribution of this paper is the use of free and open-source software (FOSS) tools for data analysis and data collection. In Brazil, around 70% of research funding is provided by government agencies (McManus & Neves, 2021), and federal investment in science has been in decline (Angelo, 2019; Rodrigues, 2021). Low-income and upper-middle-income countries still lack access to software and researchers overwhelmingly still rely on proprietary software (Vermeir et al., 2018). FOSS can be a helpful tool for academic and commercial research in low-income countries, as it helps investigators to collect data, handle databases, and organize and analyze complex digital information with software technology that is freely available to everyone. The adapted RLWM task applied in the current study, for instance, was designed using OpenSesame, easy-to-work free software with an intuitive interface that enables the construction of a myriad of behavioral tasks (Mathôt et al., 2012). Visual and auditory stimuli are easy to manipulate in this software, and additional manipulations not available in the user interface can be included using basic Python programming skills. FOSS was also used for statistical analyses and computational modeling in this study. Statistical testing was conducted using R and its most used interface, RStudio. R is an open-source language that also has many libraries designed for numerical and statistical computation, so researchers can explore statistical models using guides provided by researchers that create data analysis libraries (e.g., Revelle & Wilt, 2019).

For our data synthesis and computational analysis, the Python language was used. Python is already used in neuroscience research to simulate experiments, congregate neuroimaging data, and process raw data using machine learning methods (Muller et al., 2015). The present study used Python to replicate previous computational analyses using other software (see Collins, 2018, for instance) and implement computational models of RL – and its interaction with the WM – that fit actual behavioral data. These achievements using Python code are especially relevant for two reasons. First, the computational analysis of behavioral data is the core of computational psychiatry, a growing research field that intertwines psychiatry and computational neuroscience to understand how the interaction between context and brain computations explains ‘abnormal’ behavior (Adams et al., 2016; Huys et al., 2021). We believe that the availability of Python code with computational models can engage many

researchers in computational psychiatry since Python is a growing programming language. Second, the application of computational analysis in this study replicated previous data collected in-person and with a longer version of the RLWM task (Collins, 2018), which suggests that the RLWMI model applied here is valid between studies, between distinct methods of data assessment, and for distinct versions of the same behavioral paradigm. This endorses the idea of computational validity (Redish et al., 2022) for the RLWMI model and reinforces the assumption made in previous publications that this model is relevant to comprehend interactions between RL and WM load (Collins, 2018; Collins & Frank, 2012; Collins et al., 2014; Master et al., 2020).

There are limitations in our study that need to be mentioned. Our supervised remote experiment has not been compared to face-to-face assessment or to asynchronous data collection methods. It will be useful to do it in the future to reveal problems or advantages of our assessment method for complex and long behavioral experiments. The Internet connection latency of participants (i.e., ping responses) was not registered, and this is recommended for tasks with reaction time outcomes. The pandemic’s context was also a limitation: Besides the fact that we replicated behavioral and computational data from other studies, it is still possible that our data were influenced by the context of social isolation. Analyses comparing subjects before and after the pandemic and comparing groups of people with different levels of social isolation would be relevant to control for the effects of isolation; however, they were not within the scope of this study. The use of FOSS tools was intentional in this study because we aimed for free solutions to our assessment method. We acknowledge that the use of FOSS in our study did not remove the need of human resources in our supervised remote testing. However, given the complexity and the duration of our assessment, we do not think that our data collection would be possible without the use of human resources and additional costs from proprietary software that would just increase our research expenses. In future studies, comparisons between FOSS and paid options can be carried out to test which software tools are best suited for online assessments involving long and complex experiments.

The computational validation and the detailed description of the procedures of the remote experiment described in this paper may be useful for researchers interested in conducting long and complex experiments online. The authors are happy to provide further information about the procedures described here to researchers interested in running experiments online. The data and the codes used for data analysis and computational modeling are also available. We hope that this paper will encourage researchers to conduct complex and long experiments online using supervised remote testing as an option.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.3758/s13428-022-01982-6>.

**Acknowledgements** We thank Dr. Anne Collins for her insightful suggestions and for providing Matlab code examples for our original in-person assessment and for the computational analysis.

**Code availability** The Python code used to organize behavioral data and apply behavioral models is available at GitHub (see Palmieri et al., 2022).

**Funding** GW was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) during the conduction of this study. LB was sponsored by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

**Data availability** Researchers interested in the scripts used for data analyses can access this information on Open Science Framework repository (OSF: [https://osf.io/xgrzf/?view\\_only=5d50232f37954dc098367e7542f870f2](https://osf.io/xgrzf/?view_only=5d50232f37954dc098367e7542f870f2)). Additional information regarding procedures is also available through contact with the main author.

## Declarations

**Conflicts of interest/competing interests** The authors declare that they have no conflicts of interest.

**Ethics approval** All procedures and instruments used in this study were approved by the Ethics Committee of Hospital de Clinicas de Porto Alegre, Brazil, and registered in *Plataforma Brasil*, the Brazilian National Research Ethics Platform (62798716.2.0000.5327), in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

**Consent to participate** Informed consent was obtained online for this study from all individual participants.

## References

- Adams, R. A., Huys, Q. J., & Roiser, J. P. (2016). Computational psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, 87(1), 53–63. <https://doi.org/10.1136/jnnp-2015-310737>
- Angelo, C. (2019). Brazil's government freezes nearly half of its science spending. *Nature*, 568, 155–156. <https://doi.org/10.1038/d41586-019-01079-9>
- Bilder, R. M., Postal, K. S., Barisa, M., Aase, D. M., Cullum, C. M., Gillaspay, S. R., Harder, L., Kanter, G., Lanca, M., Lechuga, D. M., Morgan, J. M., Most, R., Puente, A. E., Salinas, C. M., & Woodhouse, J. (2020). Inter organizational practice committee recommendations/guidance for Teleneuropsychology in response to the COVID-19 pandemic. *Archives of Clinical Neuropsychology*, 35(6), 647–659. <https://doi.org/10.1093/arclin/aca046>
- Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychology Review*, 27(2), 174–186. <https://doi.org/10.1007/s11065-017-9349-1>
- Carr, D. J., Adia, A. C., Wray, T. B., Celio, M. A., Pérez, A. E., & Monti, P. M. (2020). Using the internet to access key populations in ecological momentary assessment research: Comparing adherence, reactivity, and erratic responding across those enrolled remotely versus in-person. *Psychological Assessment*, 32(8), 768–779. <https://doi.org/10.1037/pas0000847>
- Cernich, A. N., Brenna, D. M., Barker, L. M., & Bleiberg, J. (2007). Sources of error in computerized neuropsychological assessment. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 22(Suppl 1), S39–S48. <https://doi.org/10.1016/j.acn.2006.10.004>
- Chaytor, N. S., Barbosa-Leiker, C., Germine, L. T., Fonseca, L. M., McPherson, S. M., & Tuttle, K. R. (2021). Construct validity, ecological validity and acceptance of self-administered online neuropsychological assessment in adults. *The Clinical Neuropsychologist*, 35(1), 148–164. <https://doi.org/10.1080/13854046.2020.1811893>
- Collins, A. (2018). The tortoise and the hare: Interactions between reinforcement learning and working memory. *Journal of Cognitive Neuroscience*, 30(10), 1422–1432. [https://doi.org/10.1162/jocn\\_a\\_01238](https://doi.org/10.1162/jocn_a_01238)
- Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *The European Journal of Neuroscience*, 35(7), 1024–1035. <https://doi.org/10.1111/j.1460-9568.2011.07980.x>
- Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *The Journal of Neuroscience*, 34(41), 13747–13756. <https://doi.org/10.1523/JNEUROSCI.0989-14.2014>
- Cuttler, C., LaFrance, E. M., & Stueber, A. (2021). Acute effects of high-potency cannabis flower and cannabis concentrates on everyday life memory and decision making. *Scientific Reports*, 11(1), 13784. <https://doi.org/10.1038/s41598-021-93198-5>
- Daw, N. (2011). Trial-by-trial data analysis using computational models: (Tutorial review). In *Decision making, affect, and learning: Attention and performance XXIII*. Oxford University Press. Retrieved 16 Aug. 2021, from <https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199600434.001.0001/acprof-9780199600434-chapter-001>
- Donahoe, J. W. (2017). Behavior analysis and neuroscience: Complementary disciplines. *Journal of the Experimental Analysis of Behavior*, 107(3), 301–320. <https://doi.org/10.1002/jeab.251>
- Gagné, N., & Franzen, L. (2021). How to run behavioural experiments online: Best practice suggestions for cognitive psychology and neuroscience. *PsyArXiv*, 10.31234/osf.io/nt67j.
- Geddes, M. R., O'Connell, M. E., Fisk, J. D., Gauthier, S., Camicioli, R., Ismail, Z., & Alzheimer Society of Canada Task Force on Dementia Care Best Practices for COVID-19. (2020). Remote cognitive and behavioral assessment: Report of the Alzheimer Society of Canada task force on dementia care best practices for COVID-19. *Alzheimer's & Dementia*, 12(1), e12111. <https://doi.org/10.1002/dad2.12111>
- Grieve, R., & Elliott, J. (2013). Cyberfaking: I can, so I will? Intentions to fake in online psychological testing. *Cyberpsychology, Behavior and Social Networking*, 16(5), 364–369. <https://doi.org/10.1089/cyber.2012.0271>
- Grootswagers, T. (2020). A primer on running human behavioural experiments online. *Behavior Research Methods*, 52(6), 2283–2286. <https://doi.org/10.3758/s13428-020-01395-3>
- Holmlund, T. B., Foltz, P. W., Cohen, A. S., Johansen, H. D., Sigurdson, R., Fugelli, P., Bergsager, D., Cheng, J., Bernstein, J., Rosenfeld, E., & Elvevåg, B. (2019). Moving psychological assessment out of the controlled laboratory setting: Practical challenges. *Psychological Assessment*, 31(3), 292–303. <https://doi.org/10.1037/pas0000647>
- Huys, Q., Browning, M., Paulus, M. P., & Frank, M. J. (2021). Advances in the computational understanding of mental illness. *Neuropsychopharmacology: Official publication of the American college of Neuropsychopharmacology*, 46(1), 3–19. <https://doi.org/10.1038/s41386-020-0746-4>
- Huys, Q. J., Moutoussis, M., & Williams, J. (2011). Are computational models of any use to psychiatry? *Neural networks: The official journal of the International Neural Network Society*, 24(6), 544–551. <https://doi.org/10.1016/j.neunet.2011.03.001>



- Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., & Collins, A. (2020). Disentangling the systems contributing to changes in learning during adolescence. *Developmental Cognitive Neuroscience, 41*, 100732. <https://doi.org/10.1016/j.dcn.2019.100732>
- McKinney, W., et al. (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference, 445*, 51–56. <https://doi.org/10.25080/Majora-92bf1922-00a>
- McManus, C., & Baeta Neves, A. A. (2021). Funding research in Brazil. *Scientometrics, 126*, 801–823. <https://doi.org/10.1007/s11192-020-03762-5>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Muller, E., Bednar, J. A., Diesmann, M., Gewaltig, M. O., Hines, M., & Davison, A. P. (2015). *Python in neuroscience*. *Frontiers in Neuroinformatics, 9*, 11. <https://doi.org/10.3389/fninf.2015.00011>
- Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra: Psychology, 6*(1), 17213. <https://doi.org/10.1525/collabra.17213>
- Palmieri, I., Weydman, G., & Simoes, R. A. G. (2022). Reinforcement learning and working memory models (version v0.1) [computer software]. <https://github.com/ipalmieri/rlwm>
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing <https://www.R-project.org/>
- Redish, A. D., Kepecs, A., Anderson, L. M., Calvin, O. L., Grissom, N. M., Haynos, A. F., Heilbronner, S. R., Herman, A. B., Jacob, S., Ma, S., Vilares, I., Vinogradov, S., Walters, C. J., Widge, A. S., Zick, J. L., & Zilverstand, A. (2022). Computational validity: Using computation to translate behaviours across species. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 377*(1844), 20200525. <https://doi.org/10.1098/rstb.2020.0525>
- Revelle, W., & Wilt, J. (2019). Analyzing dynamic data: A tutorial. *Personality and Individual Differences, 136*, 38–51. <https://doi.org/10.1016/j.paid.2017.08.020>
- Rodrigues, M. (2021). Scientists reel as Brazilian government backtracks on research funds. *Nature, Advance online publication*. <https://doi.org/10.1038/d41586-021-02886-9>.
- RStudio Team. (2019). *RStudio: Integrated development for R*. RStudio, Inc., Boston, MA <http://www.rstudio.com/>
- Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences, 10*(4), 251. <https://doi.org/10.3390/brainsci10040251>
- Semmelmann, K., & Weigelt, S. (2017). Online psychophysics: Reaction time effects in cognitive experiments. *Behavior Research Methods, 49*(4), 1241–1260. <https://doi.org/10.3758/s13428-016-0783-4>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd Edition). MIT press. <https://mitpress.mit.edu/books/reinforcement-learning-second-edition>
- Tailby, C., Collins, A. J., Vaughan, D. N., Abbott, D. F., O'Shea, M., Helmstaedter, C., & Jackson, G. D. (2020). Teleneuropsychology in the time of COVID-19: The experience of the Australian epilepsy project. *Seizure, 83*, 89–97. <https://doi.org/10.1016/j.seizure.2020.10.005>
- Vermeir, K., Leonelli, S., Tarik, A. S. B., Sojinu, S. O., Ocloo, A., Khan, A. I., & Benzuidenholt, L. (2018). Global access to research software - the forgotten pillar of Open Science implementation. *German National Academy of Sciences Leopoldina* <https://globalyoungacademy.net/publications/global-access-to-research-software-the-forgotten-pillar-of-open-science-implementation/>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., & SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods, 17*, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wadsworth, H. E., Galusha-Glasscock, J. M., Womack, K. B., Quiçeno, M., Weiner, M. F., Hynan, L. S., Shore, J., & Cullum, C. M. (2016). Remote neuropsychological assessment in rural American Indians with and without cognitive impairment. *Archives Of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 31*(5), 420–425. <https://doi.org/10.1093/arclin/acw030>
- Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife, 8*, e49547. <https://doi.org/10.7554/eLife.49547>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.