



TOPIC REVIEW

How to design an objective test to assess personality: Step by step

Cómo diseñar un test objetivo para evaluar la personalidad: paso a paso

Miriam Romero¹ , José Manuel Hernández López¹ , José Santacreu¹ 

¹ Universidad Autónoma de Madrid, España.

How to cite: Romero, M., Hernández, J.M., & Santacreu, J. (2024). How to design an objective test to assess personality: Step by step. *Rev. CES Psico*, 17(3), 129-158. <https://dx.doi.org/10.21615/cesp.7279>

Abstract

Objective tests allow us to study the behavior that people show, without having to resort to the analysis of the verbal descriptions they make about their own behavior. Objective tests have been widely used to study abilities, and it was not until the works of Cattell and Warburton that these instruments also began to be systematically designed to study personality. This work aims to provide researchers with the necessary knowledge to design an objective personality test. Initially, The article presents a review of the main aspects of objective tests, including their historical context and theoretical foundations. Following this review, we propose a detailed methodology to guide the design of objective personality tests. This methodology includes key steps such as defining the construct, designing the test's morphology and functionality, and conducting pilot studies to ensure reliability and validity. This structured approach ensures the development of robust and precise tools for personality assessment, thereby making a significant contribution to the field of psychology. Objective tests offer notable advantages in providing unbiased and accurate assessments of personality. This work serves as a foundational guide for researchers aiming to develop objective personality assessment tools, emphasizing the importance of methodological rigor and interdisciplinary collaboration.

Keywords: objective tests; test design; assessment; personality; interactive style.

Resumen

Los tests objetivos nos permiten estudiar el comportamiento de las personas sin tener que recurrir al análisis de las descripciones verbales que estas hacen sobre su conducta. Los test objetivos han sido ampliamente utilizados para el estudio de competencias, y no fue hasta los trabajos de Cattell y Warburton cuando estos instrumentos comenzaron a diseñarse de forma sistemática para estudiar también la personalidad. Este trabajo tiene como objetivo proporcionar a los investigadores el conocimiento necesario para diseñar un test objetivo de personalidad. En primer lugar, el artículo presenta una revisión de las principales características de los test objetivos, incluyendo su contexto histórico y fundamentos teóricos. Tras esta revisión, proponemos una metodología detallada para guiar el diseño de test objetivos de personalidad. Esta metodología incluye pasos clave como la definición del constructo, el diseño de la morfología y funcionalidad del test, y la realización de estudios piloto para asegurar la fiabilidad y validez. Este enfoque estructurado asegura el desarrollo de herramientas robustas y precisas para la evaluación de la personalidad, contribuyendo significativamente al campo de la psicología. Estos instrumentos presentan entre sus ventajas el hecho de que proporcionan evaluaciones de la personalidad precisas y objetivas. Este trabajo sirve como una guía para los investigadores que requieran diseñar test objetivos para la evaluación de la personalidad, enfatizando la importancia del rigor metodológico y la colaboración interdisciplinaria.

Palabras clave: test objetivos; diseño de test; evaluación; personalidad; estilo interactivo.

Introduction

Personality refers to idiosyncratic, consistent and stable tendencies. Traditionally, the trait model has

predominated in the study of personality. This model conceptualizes personality under the label of traits, which are internal variables that have a causal role in the behavior. Psychologists working within this model have described the structure of the personality according to a set of factors (e.g., Big Five; Norman, 1963).

On the other hand, from a behavioral approach, personality is a consistent and stable behavioral trend that does not explain nor cause behavior. Rather, it is the individual's behavior itself in a specific context (Santacreu et al., 2006). The interbehavioral theory formulated by Kantor (1924) provides the basis for studying personality in observable and objective terms. From the interbehavioral theory, what is studied is the interaction between an individual and a specific context, as well as the variables that influence that interaction. Personality reflects the synthesis of learning experiences that the person has accumulated throughout his or her life. Personality is a dispositional variable that does not directly take part in the interaction, but influences it. Ribes (1990, 2018) was one of the contemporary authors who applied the interbehavioral theory to the study of personality conceptualized as an "interactive style".

In this work, we adopt a behavioral approach to measure personality. Establishing as a starting point that personality is a dispositional variable that reflects the synthesis of the experiences an individual has accumulated throughout his or her life, we can assess personality in two ways:

- a) Obtaining information about the verbal synthesis that people make of their behavior (what people say about themselves).
- b) Obtaining information about the behavior that people show in a particular situation (what people do or choose to do in a specific situation).

In the first case, we assess personality through self-reports. In the second case, we employ objective tests. Self-reports and objective tests could be complementary methods for assessing personality. Self-reports are widely known, and their use in personality psychology is widespread. However, the same does not apply to objective tests.

The aim of this work is to introduce objective tests as a useful tool for personality assessment. To this end, we will describe objective tests and the steps required to design them. In addition, we provide supplementary material that reflects the application of these steps to the design of a fictitious self-control objective test.

What is an objective test?

Objective tests are instruments to assess people's behavior in order to infer a psychological variable (Anastasi, 1954; Cattell, 1957; Scheier, 1958). These instruments have been widely used to measure variables such as spatial ability, verbal comprehension capacity or processing speed. In personality, the most employed instruments have been self-reports. However, the pioneering work of Cattell and Warburton (1967) showed that objective tests could also be designed to measure personality variables.

Self-reports present questions or descriptions to which the assessed must respond, generally with a scale of agreement. These instruments require people to evaluate and describe their own behavior. Contrary to self-reports, objective tests are not based on verbal descriptions. Objective tests are based on the sequence of observable behaviors of the people being assessed (Anastasi, 1954; Cattell, 1957, 1958; Ortner & Proyer, 2015; Scheier, 1958). According to Cattell's terminology (e.g., Cattell, 1958), objective tests allow us to obtain T-data (test-data: observable behaviors that people show in specific situations), meanwhile self-reports allow us to obtain Q-data (questionnaire-data: verbal descriptions of the behavior).

The objectivity of an instrument can be studied from two points of view. On the one hand, we can consider the point of view of the researchers that carry on the assessment. On the other hand, we can consider the point of

view of the people who are being assessed. Self-reports and objective tests are objective from the point of view of the researchers that perform the assessment, since the scores they obtain do not require a subjective interpretation (the assessors do not have to interpret the results but assign scores according to previously established criteria). Different researchers will obtain the same score. However, this is not the case if we consider the point of view of the person being assessed. As self-reports are based on verbal descriptions, when people complete them, a degree of subjectivity is sometimes inevitable. Instead, as objective tests record the sequence of responses that people show when they face a situation, the subjectivity is easily avoided. Therefore, objective tests present a double objectivity. Neither the responses of those evaluated nor the scores obtained by the researchers are affected by subjectivity (Cattell & Warburton, 1967; Ortner & Proyer, 2015).

The first objective tests designed by Cattell and Warburton (1967) had a mechanical or pencil and paper format. These tests have been called first generation objective tests (Ortner & Proyer, 2015). Later, thanks to computer advances, objective tests began to be developed in a computerized format. These are called second generation objective tests.

The following are some examples of second generation objective tests:

- Betting Dice test (Rubio & Santacreu, 1998): This test measures risk tendency. It presents two dice rolling in the screen. Participants have to make a bet about the result of adding both dice, choosing between four options: a result greater than 4, greater than 7, greater than 9 or equal to 12. The probability of obtaining the first option is greater than the second option and so on (although the expected value is identical for the four options). The lower the probability of obtaining the result, the riskier the bet.
- Balloon Analogue Risk Test (Lejuez et al., 2002): This test also measures risk tendency. The screen presents a balloon that has to be inflated. When participants inflate the balloon, they receive coins. They can choose when to stop inflating the balloon and therefore, when to collect the coins. If they inflate the balloon too much, it explodes and they do not receive the accumulated coins.
- Conscientiousness test (Hernández et al., 2003): This test presents a set of different types of figures organized in rows and columns. The figures represent different types of trees. A model tree is shown to the participants, and they are asked to select all the figures that are similar to the model. The test records the pattern of scores, obtaining a measure of conscientiousness.
- Multidimensional Objective Interest Battery (Proyer, 2006). This battery includes an objective test to assess vocational interests. For example, in one of the subtests, participants are asked to distribute 100,000 euros to a set of associations. The amount of money allocated to each association is considered a measure of interest.
- Mastery Performance-Goal Orientation Test (Romero et al., 2020): This test assesses goal orientation: mastery- and performance-orientation. It presents a matrix with different figures that belong to different categories. By clicking on the figures, the participants get points. One of the categories offers a higher number of points (optimal figures). Participants can obtain a high number of points either by learning which category is optimal or by clicking on a large number of different figures. People who try to get points by restricting the clicks to the optimal category are classified as mastery-oriented. People who click on a large number of figures without discriminating between categories are classified as performance-oriented.

Despite the potential usefulness of objective tests, they have hardly been used in personality literature. This underutilization can be attributed to inherent limitations and findings from previous research that compare them with the results of self-reports (see Dang et al., 2020 for a review).

On the one hand, a main reason that may have contributed to the lack of widespread use of objective tests is that their design involves more costs than the design of self-reports, which is a significant limitation. As we will elaborate later, the development process is intricate and involves defining the construct, designing a mechanical or computerized way of recording observable data, conducting pilot testing, readjusting the design, etc., to ensure reliability and validity. These steps are not only time-consuming but also incur substantial costs, which can be a barrier to the development of new objective tests.

Moreover, second-generation objective tests typically require computerized administration, necessitating collaboration with computer technicians for programming. Although technological advancements have made the design of these tests more feasible, they still demand ongoing maintenance and updates to remain functional and relevant amid rapidly evolving digital platforms. Hence, forming a multidisciplinary team is essential and beneficial.

In addition, and related to the above, the practical administration of objective tests requires equipment with specific technical requirements, which contrasts with the ease of administering paper-and-pencil questionnaires. This need can limit the accessibility and applicability of objective tests in various contexts.

On the other hand, an objective test often focuses narrowly on a specific personality variable, and to capture other interacting personality dimensions, it is necessary to use different objective tests. This contrasts with the ease of using a questionnaire, which can measure different personality variables through a single instrument and is potentially less time-consuming.

Finally, we believe that one of the key reasons contributing to the underuse of objective tests is that, in the research conducted by Cattell, T-data did not correlate with Q-data (e.g., Ortner & Proyer, 2015; Smith, 1988). Cattell and his team aimed to employ multiple methods to assess the same variable. The lack of correlation between self-reports and objective tests generated inconsistencies regarding their theory, which motivated personality-trait researchers to employ self-reports to the detriment of objective tests. However, the lack of correlation between T and Q data is a fact that should be expected. What we say and what we do does not necessarily coincide. For example, someone can say that he or she is conscientious. However, when observing their behavior, we may see that their desk is messy or that they solve crossword puzzles in a non-systematic way. This person may not have intended to lie. Rather, the subjectivity of language contributes to the lack of correspondence between what we say and what we do. In addition, social desirability or acquiescence are biases that can influence people to distort their descriptions of themselves. Although the results obtained by Cattell's team seem to have generated a lack of confidence in the objective tests, it is important to note that the absence of correlation between what we say and what we do does not mean that objective tests or self-reports are not valid. Objective tests and self-reports measure different aspects of the same variable (Ortner & Proyer, 2015).

We believe it is possible to take up Cattell's proposal again. The contributions of this author and his collaborators, as well as the personality behavioral approach, will provide psychologists with the basis to design objective tests to measure personality. Research can benefit from employing objective tests, as they provide complementary data to those used by other instruments. The design of these instruments requires interdisciplinary work, but this is enriching and can advance psychological science.

In the following sections, we will describe the aspects that we consider sufficient and necessary for a psychologist to start designing an objective personality test.

Designing an objective test

General requirements

Before describing step by step how to design an objective test, it is important to know the general requirements

in order to assess personality through this type of instrument. Let us begin by briefly describing the behavioral theory of personality that will be the basis of the objective test design. We will integrate the contributions of Cattell but, unlike his group and the vast majority of psychologists attempted, we will not identify personality traits, we will identify behavioral patterns or interactive styles (Santacreu et al., 2002).

Behavioral theory has an obvious interest in observable behavior and aims to explain behavior based on context contingencies. However, recently, this theory has also been interested in the role of people's history, synthesized throughout the life of the individual. The following is a summary of the fundamentals that allow us to study personality.

Functional relationships between individual's behavior and specific context events explain the change of the behavior in that context. In other words, in a particular moment and situation, individual behavior (B) is a function (f) of the context contingencies (C). People are different because of genetic, epigenetic, or ontogenetic experiences. We postulate that individuals constantly synthesize their experiences and learnings in order to act (respond or decide) as quickly as possible (as mentioned before, the synthesis of their history is expressed both as behavioral trends and as verbally expressed rules of behavior). When we study the differences in the behavior between individuals, we consider that the behavior (B) is a function of the characteristics of the context contingencies (C), but depends on the dispositional factors of the individual (I).

$$B = f(I, C)$$

Cattell and his group identified three different areas that allow classifying a person's history: abilities, personality, and motivation (e.g., Cattell et al., 1972). Substituting the individual (I) for each of his areas, the following formula reflects that the behavior (B) is a function (f) of the mentioned categories of the individual [Motivation (M), Abilities (A) and Personality (P)], and the contingencies of the context (C).

$$B = f[(M, A, P), C]$$

To study the role of the characteristics of the individual in the explanation of behavior, according to the previously described formula, it is essential that the contingencies of a specific context remain stable. Furthermore, the context must not influence the participant's behavior. In this way, only the characteristics of each person will influence behavior. Therefore, the test that will measure the variables of each individual has to present an identical context for all the participants and must not report the result or consequences of the behavior that we aim to measure.

$$B = f(M, A, P)$$

In the same way, to study individual's personality, motivation and skills must not be influencing the behavior that we are going to measure. This means that, first, in order to study personality through a test; all participants should show motivation to complete it. If they are not motivated (if they do not perform any response), we will not be able to measure the variable of interest. Secondly, the individual must have the necessary skills to complete the test (for example, he or she should know how to use the mouse in a computerized test). Otherwise, ability will contaminate the results. If the behaviors of people who complete the test are not differently influenced by motivation or abilities, then we could measure the personality through these behaviors.

$$B = f(P)$$

We should remember that, in all cases, when administering an objective personality test we will be measuring the observable behavior while the individual is facing (doing, choosing or solving) a task in a specific context.

Applying all of the above, the standards that an objective test must meet in order to measure a personality

variable are the followings:

1. The test should allow us to measure personality based on observable responses. It will be necessary to determine which data the test should record in order to infer the personality variable. Generally, objective tests record data such as the sequence of responses, the frequency, latency, the duration of a response, etc. (Cattell & Warburton, 1967; Ribes, 2018; Santacreu et al., 2002; Scheier, 1958).
2. The test instructions should be clear, simple and concise. They should not allow participants to guess which variable the test measures, that is, the variable should be masked. If so, it will be more difficult for participants to distort their answers in a certain direction (Cattell & Warburton, 1967; Santacreu et al., 2002, 2006).
3. Related to the above, the test should not offer feedback on the personality variable being measured (Ribes, 2018; Santacreu et al., 2002, 2006).
4. The task has to be simple enough so that the competencies do not influence the execution of the participants (Cattell & Warburton, 1967; Santacreu et al., 2002). If a competence is needed to perform the test (e.g., spatial orientation), we could only assess personality on the participants that are competent in it.
5. The test should show adequate psychometric properties referred to observable variables: reliability and validity.

The following requirements should be verified after employing objective personality tests to assess personality of an individual.

1. Participants must have shown a minimum level of motivation when facing the task proposed. We should establish a motivation criterion that determines the minimum number of responses that a participant must show so that we can assess his or her personality (Santacreu et al., 2002).
2. Participants must have shown an undoubted level of competence when performing the task. In some tests, we should establish a competence criterion. If a participant does not meet the criterion, we could not assess his or her personality (Santacreu et al., 2002).
3. Participants must have shown consistency throughout the trials of the test (note that here we refer to individual's consistency, not to the consistency of the group data). The assumption of individual's consistency has generated an intense debate in personality psychology (see Hernández et al., 2006). If people do not show consistency, we cannot make predictions about their behavior. Therefore, it will be important to check the degree of consistency they show in the test.

Once we have described the general requirements, we will move on to explain the specific steps needed to design an objective test. In order to design an objective test, it is necessary to prepare a test description document. This document has different parts. First, it should describe the variable that we are going to measure. Secondly, it has to describe the test as it is presented to the participants. Third, it has to describe the variables that the researcher configures before administering the test (in order to change their parameters if necessary) and the variables that the researcher obtains after administering the test. This document should be prepared to be comprehensible not only to psychologists but also to the computer technicians responsible for programming the test. As mentioned, a multidisciplinary team will be both essential and beneficial. [Fig. 1](#) summarizes the sections that this document should have. In the supplementary material ([Appendix 1](#)), readers can find an example of a fictitious self-control test.

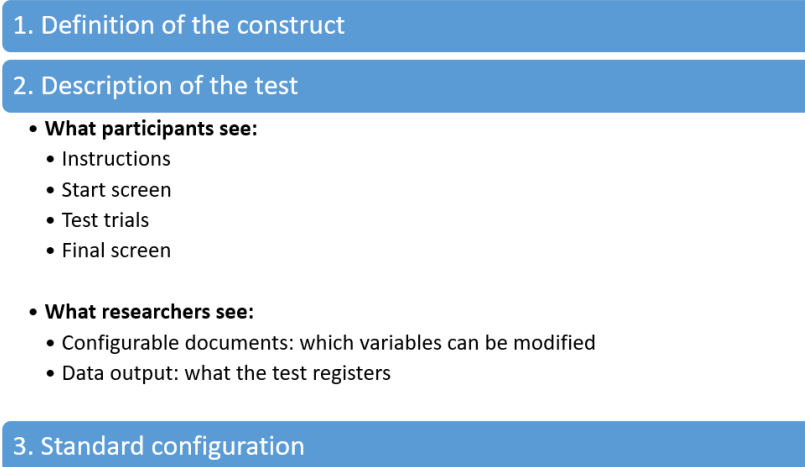


Fig. 1. Sections of the test description document of any objective test.

Steps for designing an objective test

1. Choose and understand the personality variable

First, we need to choose and understand the personality variable that is going to be measured. It is necessary to carry out a literature research on the construct to know how the most important theories that have studied it, how it is defined, in what context it is important and how it has been measured. The test description document should start with a summary of this information taking into account, first of all, the basic research on the personality variable that we intend to measure. If we do not do so, it will be impossible to design a test to measure it.

2. Think beforehand how the variable could be recorded behaviorally

The variable should be measured based on observable responses, a task that presents considerable challenges. At this early stage, it is essential to develop a general idea about how the variable could be behaviorally measured. This involves brainstorming and identifying a range of behaviors or responses that accurately reflect the underlying construct we aim to assess. Later on, and once we have thought about the morphology and functionality of the test, it will be when we will decide which specific variables our test will record. For example, to measure risk tendency, we must consider whether the participant chooses risky options, but depending on the morphology of the test, the variables to be recorded can vary. The Balloon Analogue Risk Test (Lejuez et al., 2002) records how many times the participant presses the pump to inflate a balloon. The more the participant presses, the higher his or her risk-tendency. The Betting Dice test (Rubio & Santacreu, 1998) registers in which of the four options presented in each trial the participants click. The test assigns 1 point to the 1st option (the lower risky bet), 2 points to the 2nd, option and so on. The mean score of the test trials is the score of each participant. The higher score, the higher the risk-tendency of the participant. Both tests measure risk tendency in observable terms, but the specific variables they record are slightly different. As mentioned before, objective tests usually record the latency, frequency and/or the duration of the response, the correct responses and the errors, the sequence executed, etc.

3. Design the test morphology and functionality

This step requires researchers to be creative and thoughtful about how the test should operate to measure the variable of interest effectively. In designing the test morphology and functionality, it is necessary to consider how the test will function in practice, ensuring that the measurement of the variable is accurate and reliable. In addition, to measure personality, the variable should be masked, that is, the participants should not be able to easily guess which variable is being measured.

Some objective tests are designed as a game or a competitive task, which helps to motivate participants to take it. For example, the Balloon Analogue Risk Test (Lejuez et al., 2002) can be presented to the participants as a game in which the balloon has to be inflated to obtain coins. Sometimes it will be difficult to present the test as a playful task, but if we present it with an attractive graphic design, we can also favor the motivation of the participants to complete it.

As we have announced, this step is likely one of the most challenging and may take considerable time—potentially weeks or even months—to complete. We recommend that researchers review already published objective tests to have a reference for how to obtain observable recordings of variables. Additionally, it may be useful to review and analyze existing games or video games available. Sometimes, even if they were not designed for this purpose, some existing games contain a basis for measuring certain psychological variables. For instance, regarding risk tendency, researchers might ask themselves, considering the definition of the variable, is there any board game that can measure this variable? What modifications could be made to measure risk tendency? It can also be useful to think about everyday activities or sports, as they can sometimes serve as a basis for measuring variables. Continuing with the example of risk tendency, a person who practices climbing might ask how they could design a situation to measure risk tendency in that context. We also recommend that researchers brainstorm various designs and discuss them with their colleagues. Collaboration and the exchange of ideas with other professionals can foster creativity and lead to innovative solutions.

It will be necessary to determine how the task will work (for example, will there be a time limit? If so, what will happen if the participant does not respond before the time limit?). We will also have to consider whether the test will have different phases and series. A phase is a group of trials with similar characteristics. For example, in a persistence test, there may be a first learning phase of, for example, 12 trials and a second extinction phase of, for example, 10 trials. We may assess persistence as the number of responses in the extinction phase (second phase). The trials of a specific phase that share one or more characteristics can also be grouped as a series.

In a parallel way, we will have to think about what the participant will see and what the researcher will see and describe it in the test description document. In the following, we describe what needs to be included when describing these sections (see also [Appendix 1](#) for an example).

What participants see: instructions, test trials and final screen

When the participants complete an objective test, they will have to read the instructions that indicate how the task works. Later, they will face the test trials. Finally, they will see a screen that indicates the end of the test. We will describe these parts in detail:

Instructions. The instructions should describe clear and briefly the purpose of the task. Researchers can also include in this section images or a video that shows how the task works.

Instructions should not reveal which variable is being assessed, and they should not clearly direct the behavior. It is important that the participants execute the responses in a natural way.

In the test description document, we should describe how many screens are needed to show the instructions, how are these screens and the text (or visual sources) that will be presented.

The instructions will be configurable by the researcher (see section “What the researchers see: configurable variables and data output” below).

Test trials. In these trials, the participant will perform the test. Again, the researchers will be able to configure the number of trials, the text and the feedback (see section “What the researchers see: configurable variables and data output” below). Sometimes, it will be adequate to configure initial trials (usually called training trials),

so that participants become familiar with the task.

As mentioned before, the test should not offer feedback on the variable that it assesses. The test can provide feedback on aspects other than the main variable. In addition, in some tests it will be appropriate to offer feedback on the remaining time to maintain participant motivation high.

Final screen. Participants must know that the test has been completed. Therefore, in the test description document it will be necessary to indicate how this final screen looks like. Generally, it is a screen with a simple message ("The test has finished") and a button under the text indicating "Exit".

What researchers see: configurable variables and data output

The researchers will have access to files of the computer program where they can configure the instruction screens and the variables of the test. In addition, they will have access to files with the data output.

When designing the objective test, we have to make it clear in the test description document which elements the researchers will be able to configure. In addition, when designing the test, we should indicate in the test description document which variables will be visible in the data output once the researchers have administered the test to a participant.

Data output: The objective test must provide data per trial, per series, per phase and overall test data (generally, total scores and means). The data will only be available to the researchers.

First, the data output should show if the participant has fulfilled motivation and competence criteria. The motivation criterion usually requires a minimum number of responses. The competence criterion usually requires a minimum number of hits or a maximum number of errors. If the participants do not fulfill the competence criteria, participants cannot be assessed by the test.

Only if the participant has fulfilled motivation and competence criteria, we can assess their personality.

4. Design standard instructions and a standard configuration

Up to this point, we will have already designed the morphology and functionality of the test and will have considered which variables the experimenter should configure. At this stage, we will have to design the standard instructions and a standard configuration that we consider adequate to measure the variable of interest. This will be the base configuration for future pilot studies and may be adjusted if required. The standard instructions and the standard configuration should be included in the test description document.

5. Conduct pilot tests and make adjustments

Once we have a preliminary version of the test and a standard configuration programmed, it will be necessary to carry out pilot studies with samples of 10 or 20 participants. We must check that the software program runs without errors and that the data output is correctly registered. In addition, we should check if the participants fulfill the motivation and competence criteria to assess the variable of interest.

[Fig. 2](#) summarizes what we need to check in pilot studies. These studies will allow us to adjust the instructions and the configuration of the test in order to run a version of the test that allows us to measure the personality variable adequately. On some occasions, the participants do not properly understand the instructions that we design in the first moment, so it will be necessary to change them. In addition, sometimes the configuration that we consider as a starting point is not proper. For example, we may consider a specific time limit but, when administering the test, we could see that it is insufficient for the participants to complete the test adequately.

Before administering the test

- The variable is sufficiently and properly described
- The elements of the test that the researchers can configure are properly described
- The variable being assessed is masked
- The standard instructions of the test are brief and clear
- The standard instructions of the test do not direct participants' behavior
- The standard configuration of the test is clear and based on the theoretical descriptions of the variable
- The feedback provided in the trials does not directly offer information about the variable being assessed
- The criteria to determine if a participant is motivated is properly defined
- The criteria to determine if a participant is competent is properly defined

After administering the test

- The test does not show programming errors
- The test calculates the data correctly
- The participant shows acceptable motivation to perform the test.
- The participant shows a minimum competence to perform the test
- Variability is observed in the results of a sample of participants

Fig. 2. Checklist for the pilot studies to check the functioning of an objective test

6. Study the psychometric properties of the final version

Once we have constructed the final version of the test, we should administer it to larger samples to study its psychometric properties. The test should show reliability and validity (see, for example, DeVellis, 2016). We should demonstrate that the final version of the test allows us to measure the variable as described in the theoretical approaches (see Cronbach & Meehl; 1955). To obtain evidence of validity, we should administer our test together with another test in which the task presents a different morphology but a similar functionality (that is, that measures the same variable but presents a different-looking task). We should also study construct validity by studying whether the test shows a significant relation with variables with which, according to the theory, it should be associated. It is important to emphasize again that the relation between the test score and the behavior observed in real situations cannot be high. In real situations, the behavior of the individual depends on their skills, motivation and personality, but also on the contingencies of the situation. On the contrary, in a test, the task provides no feedback of execution, and the behavior is not therefore a function of the contingencies of the task.

7. Update the test when necessary

Although we have produced a final version of the test, this does not mean that it is invariable. Advances in theory may mean that it is necessary to review whether the test continues to fit the theoretical approaches. In addition, in some cases it may be appropriate to try out different configurations of the test to determine which of them allows us to obtain a better measurement of the variable.

On the other hand, we should not lose sight of the possible updates required from the technological point of view. Technology can advance rapidly, and this may mean that tests need to be updated so that it can be administered in the most recent browsers and computers.

Conclusions

In this work, we have presented what we consider fundamental to design an objective personality test. We have taken up Cattell's contribution, but establishing the behavioral theory of personality as a starting point.

To design an objective test is necessary to understand the variable we want to measure and to think in detail about how to measure it in observable terms. At this point, the experimental contributions of basic psychology

are fundamental. Ultimately, the variable will be measured from the data obtained by a computer, which means that we have to apply our psychological knowledge to numerical variables that are more characteristic of computer science.

Designing an objective test takes time and effort, presenting a significant challenge. However, the effort is worth it. The level of precision achieved by using this type of instrument is enormous (objective tests allow to measure several responses, the time in which each one is made, its duration, etc.). In the research field, the use of objective tests will allow us to expand the conceptual analysis of personality variables. In the research but also in the applied field, objective tests will provide us with information free of subjectivity and bias. It is necessary to assemble a multidisciplinary team for this endeavor, but this approach is enriching and beneficial, as it brings diverse perspectives and expertise that can lead to more robust and innovative solutions.

We believe that, if objective tests to measure personality are implemented in the scientific community (as has happened with competency tests) the scientific study of personality might lead to a paradigm shift in Psychology. Researchers who study the behavior of organisms in extremely controlled situations will have to accept the importance of the ontogenetic history of individuals, especially humans. People are entities with a great capacity for learning. Like all living beings, humans grow and transform as individuals, accumulating what they have learned and synthesizing it in such a way that they can respond quickly if necessary in any context. It mobilizes their systems, organs, tissues and cells as a whole, to respond quickly (although not always successfully) to change in the context in which they live. Therefore, it is important to propose a model that explains how the individuals manage to reorganize what they are learning, and how they use their learnings to act effectively at another time.

It is at this point that the objective and accurate measurement of skills, motives, personality, competencies and knowledge is radically important. Without precision in the measurements, we cannot imagine models of the functioning of people in their relation with the context in which they live. Currently, we can only observe with some precision the changes that occur in an individual in a specific context, in the short term.

In our not so humble opinion, accurate and reliable measurements can aid progress in understanding human behavior. We hope that the rationale provided in this article will serve other researchers interested in designing objective tests to assess personality variables, which would make an important theoretical and methodological contribution to the literature.

References

- Anastasi, A. (1954). *Psychological testing*. Mcmillan.
- Ashe, M. L., & Wilson, S. J. (2020). A Brief Review of Choice Bundling: A Strategy to Reduce Delay Discounting and Bolster Self-Control. *Addictive Behaviors Reports*, 100262. <https://doi.org/10.1016/j.abrep.2020.100262>
- Cattell, R. B. (1957). *Personality and motivation structure and measurement*. World Book.
- Cattell, R. B. (1958). What is "objective" in "objective personality tests?" *Journal of Counseling Psychology*, 5(4), 285-289. <https://doi.org/10.1037/h0046268>
- Cattell, R. B. & Kline, P. (1977). *The Scientific Analysis of Personality and Motivation*. Academic Press.
- Cattell, R. B., & Warburton, F. W. (1967). *Objective Personality and Motivation Tests: A theoretical introduction and practical compendium*. University of Illinois Press.
- Cattell, R. B., Barton, K., & Dielman, T. E. (1972). Prediction of School Achievement from Motivation, Personality, and Ability Measures. *Psychological Reports*, 30(1), 35-43. <https://doi.org/10.2466/pr0.1972.30.1.35>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <https://doi.org/10.1037%2Fh0040957>
- Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, 24, 267-269. <https://doi.org/10.31234/osf.io/v796c>
- DeVellis, R. F. (2016). *Scale development: Theory and Applications* (Vol. 26). Sage publications.
- Hernández, J. M., Rubio, V. J., Revuelta, J., & Santacreu, J. (2006). A Procedure for Estimating Intrasubject

- Behavior Consistency. *Educational and Psychological Measurement*, 66(3), 417-434. <https://doi.org/10.1177/0013164405275667>
- Hernández, J. M., Sánchez-Balmisa, C., Madrid, B., & Santacreu, J. (2003). La evaluación objetiva de la minuciosidad. Diseño de una prueba conductual. *Análisis y modificación de conducta*, 29, 455-477. <https://dialnet.unirioja.es/descarga/articulo/792233.pdf>
- Kantor, J.R. (1924). *Principles of psychology, vol. 1*. The Principia Press.
- Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk-taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8(2), 75-84. <https://doi.org/10.1037/1076-898X.8.2.75>
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66(6), 574-583. <https://doi.org/10.1037%2Fh0040291>
- Ortner, T. M., & Proyer, R. T. (2015). Objective Personality Tests. In T. M. Ortner y F. J. R. Van de Vijver (Eds.), *Behavior-based assessment in psychology* (pp. 133-149). Hogrefe.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 17(1), 15-22. <https://doi.org/10.1901/jeab.1972.17-15>
- Ribes, E. (1990). *Problemas Conceptuales en el Análisis del Comportamiento Humano [Conceptual Problems in the Analysis of Human Behavior]*. Trillas
- Ribes, E. (2018). *El Estudio Científico de la Conducta Individual: Una Introducción a la Teoría de la Psicología [The Scientific Study of Individual Behavior: An Introduction to the Psychology Theory]*. El Manual Moderno.
- Romero, M., Hernández, J. M., Juola, J. F., Casadevante, C., & Santacreu, J. (2020). Goal Orientation Test: An Objective Behavioral Test. *Psychological Reports*, 123(4), 1425-1451. <https://doi.org/10.1177/0033294119845847>
- Rubio, V. J., & Santacreu, J. (1998). *The Betting Dice Test*. Registration number, R.P.O.: M-70573.
- Santacreu, J. (2013). Behavioral variability and consistency: Experimental bases for a psychological theory of personality. *Escritos de psicología*, 6(1), 20-27. <https://dx.doi.org/10.5231/psy.writ.2013.0802>
- Santacreu, J., Hernández, J.M., Adarraga, P., & Márquez, M.O. (2002). *La Personalidad en el Marco de una Teoría del Comportamiento Humano [Personality in the Framework of a Theory of Human Behavior]*. Ediciones Pirámide.
- Santacreu, J., Rubio, V. J., & Hernández, J. M. (2006). The objective assessment of personality: Cattells's T-data revisited and more. *Psychology Science*, 48(1), 53-68. <https://psycnet.apa.org/record/2006-07428-004>
- Scheier, I. H. (1958). What is an "objective" test?. *Psychological Reports*, 4(1), 147-157. <https://doi.org/10.2466/pr0.1958.4.g.147>
- Smith, B.D. (1988). Personality: Multivariate systems theory and research. In J.R. Nesselroade & R.B. Cattell (Eds.). *Handbook of multivariate experimental psychology* (pp. 687-736). Plenum.

Appendix 1

Human Calculator Test Test Description Document

1. Definition of the construct

Human Calculator Test (HCT) is an objective test that assesses self-control. Self-control refers to decisions in which a high-value but long-term reward is chosen over a low-value and short-term reward. For example, if a professor who attends a congress chooses to go over the notes of the presentation that he or she has the next day instead of sightseeing around the city, we would say that he or she shows self-control. If a university student decides to study for an exam instead of playing a videogame, we would say that he or she shows self-control (Ashe & Wilson, 2020; Brigham, 1982; Rachlin, 1974; Rachlin & Green, 1972).

Self-control can be assessed by presenting two options that

- require different time intervals and.
- provide different amount of reward.

The value of the time interval is established in terms of time (e.g. seconds, minutes...). The value of the reward can be set in monetary or non-monetary terms (e.g. points). For example, Option A will provide 10 points and will require 15 seconds to complete. Option B will provide 100 points and will require 60 seconds to complete.

Those who choose the option that require a higher time interval and provide a higher amount of reward (Option B) on more occasions are the ones who will be classified as more self-controlled.

Experiments designed to study self-control usually employ procedures in which the time interval or the reinforcement offered by an option vary along trials (Ashe & Wilson, 2020; Rachlin & Green, 1972). The HCT will allow researchers to employ these types of procedure, as they will configure the values (time interval and reward) of the initial and the following trials.

General description of the test

The HCT is a self-control test that masks that assessment variable by presenting several arithmetic operations to solve. Therefore, the test-taker will think that the goal of the task is to correctly solve the operations to achieve points. Nevertheless, the participant's level of self-control is being assessed by making him or her to choose between different options of arithmetic operations to solve and analyzing his or her choices.

The test will present several screens with instructions before the test screens start. At the end, the test will show a final screen. The test screens are called trials. The researcher will be able to configure all of these screens.

The test trials will be those in which we will assess self-control. In these screens, participants will have to choose between two options with rewards of different value (points) and different time interval (seconds). As explained, to mask the self-control assessment, the test will be presented to the participants as a task in which they have to perform arithmetic operations.

- One of the options involves solving arithmetic operations that do not require much time (one-digit sums). This option has a low reward associated.
- Another option involves solving arithmetic operations that require more time (three-digit sums). This option has a high reward associated.

It is important to note that the interval time and the value of each option is fixed by the researcher. Therefore, the participant's ability to solve the operations is trivial and the correct or wrong answers are not relevant. Masking the variable will make it difficult for the participant to easily guess what the test measures. If the participants knew which variable the test is measuring, they could distort their answers, which would prevent us from measuring their natural tendency (e.g. Cattell & Warburton, 1967; Santacreu & Hernández, 2017).

Once the participants choose one option, they will have to complete sums in the given time (they will have to complete at least one sum correctly). Once the time has elapsed, a new trial will begin. In this new trial, the time interval or the value of the reward may be different to the one in previous trial (the participants will not know it, but the time interval or the value of the reward will change depending the configuration established by the researcher). After each trial, participants will receive the points indicated by the chosen option. If a participant completes more sums than in the second trial of the series, he or she will receive extra points (configurable by the researcher). We consider that, after the forced trials and the first trial, the participant will manage to complete an acceptable number of sums, and therefore we establish the second trial as the reference for the bonus.

As mentioned before, the test will allow the researcher to configure the characteristics of the options presented: the number of trials (option screens), the value of the rewards, the time intervals and the amount of time interval or reward that is modified along trials. The researchers can establish the initial values of the two options of a series and the values of the following trials of the series.

The test will register which options the participants chose in each trial. In addition, the test calculates the reinforcement rate (points/time interval) of each option. Based on the reinforcement rate, the test will provide a self-control measure. The self-control variable is conceived as the amount of B choices, that is, the number of times that the participant opts for the self-control option, minus the number of A choices or non self-control option. Nevertheless, since the points and the time interval may vary along trials, the reinforcement rate that the option offered at the time of the election is taken into account. Reinforcement rate is calculated as the points offered by the option chosen / time interval of the option chosen. Therefore, the self-control variable sums the reinforcement rates of B choices and rests the reinforcement rates of A choices. Participants that choose more B options will obtain a higher value in the variable. The data recorded will be available to the researchers.

At this point, it is important to mention that, to measure the self-control variable:

- the researcher should configure the trials so that choosing B option means more self-control, that is, the choice of a high-value but long-term reward, and
- the test should present several trials.

Trials that share one or more characteristics will be grouped together as a series. One series is the equivalent to a test trial, since the elections during that series is grouped and from there it is obtained a measure of self-control. We consider that there should be at least six trials per series.

Besides, the series that share one or more characteristics will be grouped together as a phase. The phases are programmed to obtain more data and will be proportionally similar. Therefore, each phase will include the same series, but the rewards will be increased or decreased proportionally in order to make the participant feel that it is a different trial. In addition, the order of the series could be modified.

In the configuration that we will propose (see "Standard configuration" section), the first three trials are not taken into account in the self-control variable calculus since they are established so that the participant understands, based on the feedback, the number of points offered by each option and the time needed to complete it. In addition, as the values (time interval and number of points offered) during the series will change

during the series, the initial trials will let participants to notice the trend of change in the values. In the standard configuration proposed, during the first trials, choosing one option is more effective than choosing the other, but this balance will be reversed during the last trials of the series. Therefore, we expect participants to begin the series by clicking on one option and switching in the middle of the series to the opposite option. As the self-control variable takes into account the reinforcement rate and is calculated based on the last three options of the series, the timing of the change in choices will determine the level of the self-control variable.

On the other hand, it is important to mention that, if the participant is not motivated to perform the test or if he or she does not show a minimum level of competence when performing the test, we could not measure personality (Santacreu & Hernández, 2017). We should establish a motivation and competence criterion and, if the participant does not fulfill them, his or her scores on the HCT will not be considered.

- Motivation criterion: the participant has not made the minimum number of clicks in each trial (3 clicks: 1 click on the option and at least 2 clicks on the numbers on the screen in which a sum should be solved; configurable).
- Competence criterion: the participant has not completed a minimum number of sums correctly in each trial (1 sum; configurable).

The general considerations of the test are the following:

- The test can be executed either from a local folder on the computer or from a link in a web browser in case it is implemented on a web server
- The test can run with different configurations
- The test will run in "full screen" mode optimized for a 1024x768 resolution.
- The interruption sequence is [Ctr+F12] + [Ctr+F11].

To describe the specific characteristics, this document will be divided between the part of the test seen by the participants and the part seen by the researchers.

What participants see: Modules

The test presents four modules that are executed sequentially:

1. Instructions
2. Start screen
3. Test
4. Final screen

What researchers see: configurable documents and data output

Configurable documents

Each module has configurable elements. The researcher will be able to modify three documents:

- a) Configuration of the instructions and the start screen (instr.json)
- b) Configuration of the test and the final screen (config.json)
- c) Set of arithmetic operations (calculator.json)

Data output

The researcher will be able to access the data recorded by the test.

2. What participants see: Modules

2.1. Module 1: Instructions

This module consists of two pages of instructions. On the first page, there are short instructions (see [Figure 1](#)).

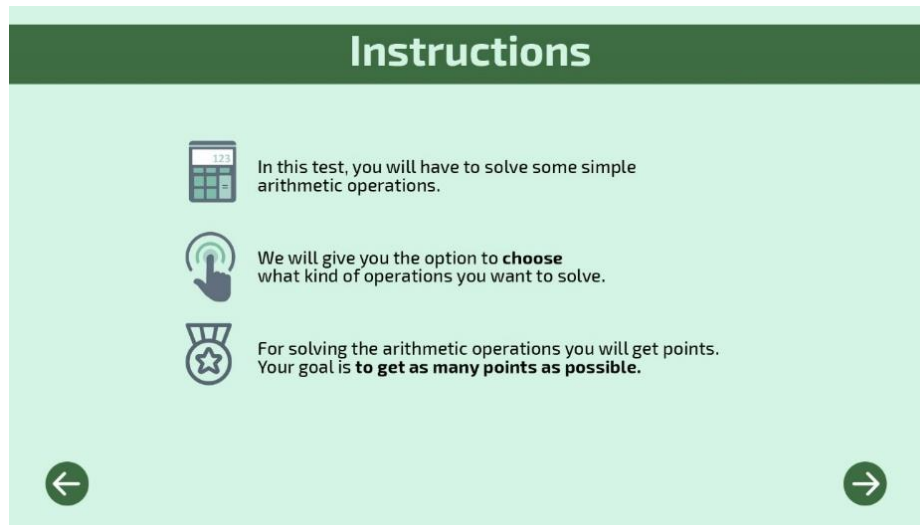


Figure 1. First instructions screen.

On the second page, the screen will show a reminder (see [Figure 2](#)).

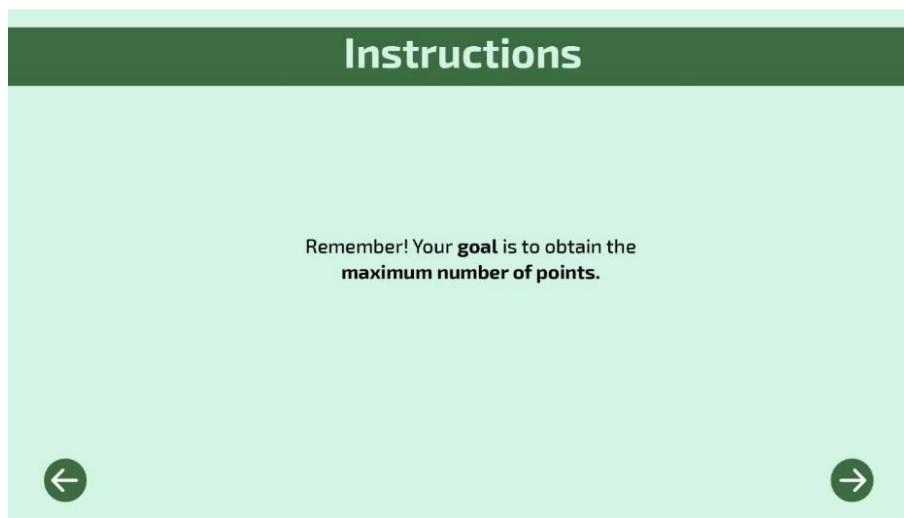


Figure 2. Second instructions screen.

2.2. Module 3: Start screen

In this module, the participant is given the opportunity to start the test or to go back and review the instructions (see [Figure 3](#)).

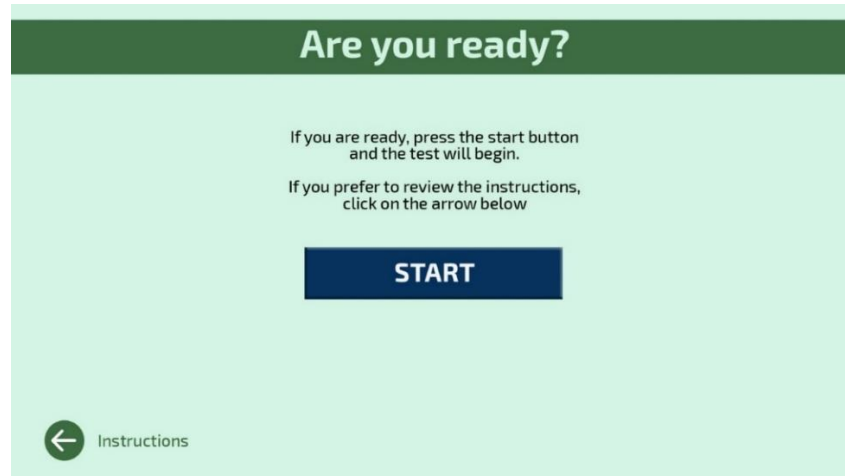


Figure 3. Start screen.

2.3. Module 3: Test

In the test trials, the participants will be presented with two options to choose. The trials will be grouped as series (a series is a set of trials that share one or more characteristics). On the other hand, series will be grouped as phases. As mentioned before, each series should have six trials or more.

The first pair of trials of each series is intended to show the participants what they will face when choosing each of the options. In the first trial, one of the options will be deactivated (e.g. Option B) so that they are forced to choose one and see what happens next. The button only presents information about option A or B and the type of tasks (e.g. one digit sums; 3 digit sums). In the second trial, the opposite option will be deactivated (e.g. Option A) so that they are forced to choose the remained one and see what it happens. After the pair of forced trials, the test trials will begin. In the test trials, both options will be available so that the participants choose one of them (see Figure 4). In the screen in which the options are presented, the remaining time of the test may be presented (configurable) in the left panel. The remaining time will be based on the theoretical maximum duration of the test and its function is to encourage the motivation of the participants.

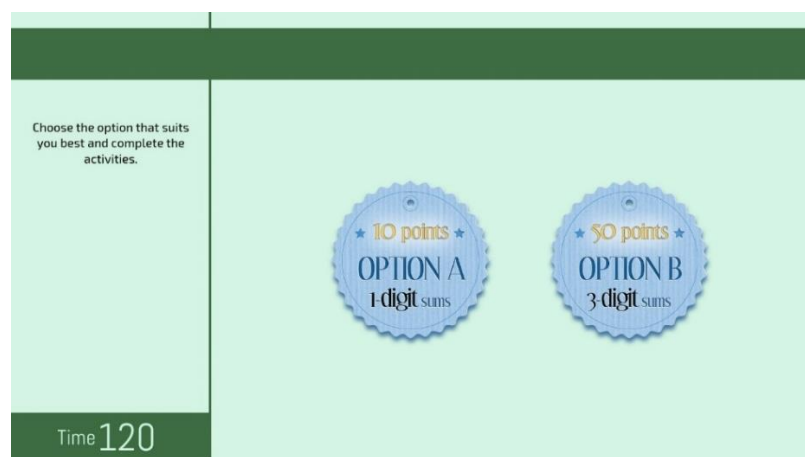


Figure 4. Test trial.

What will participants see when choosing an option? They will be presented with a screen in which they have to complete a sum. If they press Option A, they will have to solve a one-digit sum. If they press Option B, they will have to solve a three-digit sum (see Figure 5). The researcher will have previously determined a set of sums that may appear. The program will randomly choose one of those sums and will be able to detect if the numbers are

correct.

The participant must click on the numbers (grey buttons) to write the result. When clicking on a number, it is placed in the white rectangle. If the number is correct, the number remains in the white rectangle and it appears with a green border. If the number is not correct, so that the participant notices the mistake, the border of the rectangle appears in red, the buttons deactivate and the number flashes for a few moments after being deleted automatically. Once the number is deleted, the buttons activate again.

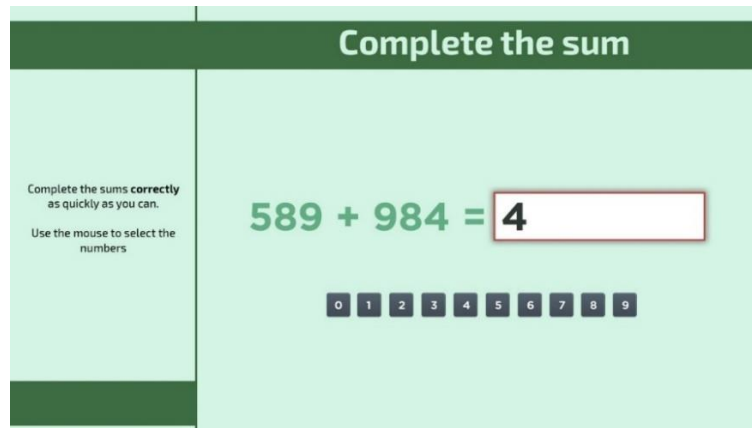


Figure 5. Test trial: sums.

The participant then must click on the numbers again. When the participant enters the correct result, the program automatically presents another sum. The researcher will also have previously determined the time interval during which new sums will appear. If that time ends, the program will present a new trial, even if the participant was still solving a sum. This implies that the time interval will be the same for all participants but not all of them will complete the same number of sums. The score achieved will vary, being slightly superior if the participant has completed more sums than in the last trial of the current sequence.

In the next trial, on the left of the screen, feedback on the points accumulated can be provided (configurable). On the right of the screen, new options are displayed. After the first pair of forced elections, the participant will have to choose the option based on the feedback received previously. The new options will not show how many points they will provide, but after completing the sums, they will receive feedback on the points received (see Figure 6).



Figure 6. Test trial: feedback.

The participant will have to choose an option and the process will be repeated: sums will appear and he or she will have to complete them.

If a certain number of seconds (configurable) have elapsed without the participant choosing an option, a reminder will be included (“Please, choose an option to continue”).

The researcher will have previously determined how the value of the options will be modified from one trial to another (series) and how many phases (set of series) the test will run. In addition, the color of the options may change between series (configurable), and the color of the background may change between phases (configurable).

Once the phases configured by the researcher are completed, the test will end, presenting the final screen.

2.4. Module 4: Final screen

The final screen is the one that lets the participant know that the test has finished (see [Figure 7](#)).

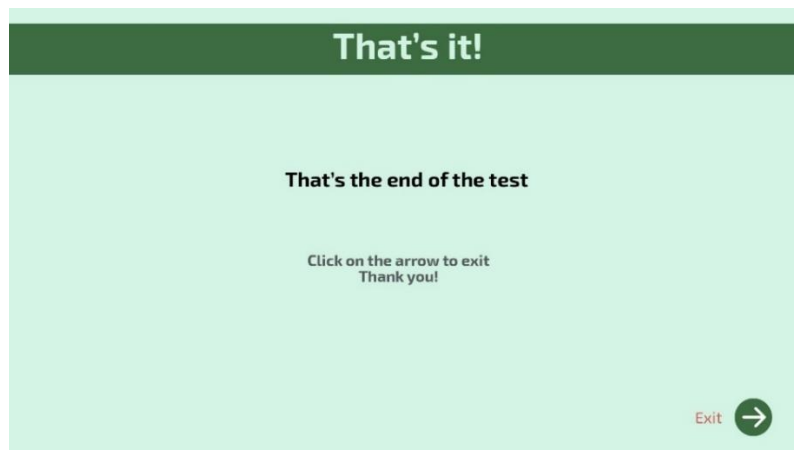


Figure 7. Final screen.

3. What researchers see: Configuration files

As mentioned above, the researcher will be able to modify three documents: the instructions (and start screen), the test configuration, and the set of sums.

3.1. Instructions and start screen

The instruction file will contain:

- the text that appears on the instruction screens and
- the text that appears on the start screen

The researcher will be able to modify the elements by accessing the file.

3.2. Set of sums

This file will contain the set of sums that can be presented. There will be more sums than can be presented in a trial.

3.3. Configuration of the test

This file will determine the functioning of the test. It will specify both the general conditions of the test and the particular conditions of each phase, series and trial. The variables that can be configured by the researcher will be presented in four sections: general, phase, series and trial.

General.

| Variable | Format | Values | Description |
|--------------------------------|--------|---|--|
| Language | X | EN, SP | ISO 639-1 code that corresponds to the language in which the test will be presented. EN = English; SP = Spanish. |
| Time_test | X | -1, 1 to N -1: no time limit | It indicates the maximum theoretical duration (seconds) of the test. This value and its progression (reduction) may be shown in the left panel on the test trials. |
| Feedback_Time_test | X | 0 = no feedback 1=feedback | Indicates if feedback of the remaining time is shown in the left panel on test trials. It will have a decreasing progression. |
| Number_phases | X | 1 to N | Purely informative variable. It informs about the number of phases of the test. |
| Number_series | X | 1 to N | Purely informative variable. It informs about the number of series of the test. |
| Number_series_phase_j | X | 1 to N | Purely informative variable. It informs about the number of series of each phase. |
| Number_trials_phase_j_series | X | 1 to N | Purely informative variable. It informs about the number of trials of each series. |
| Color_background_phases_change | X | 0 = no color change 1 = color change in order 2 = color change random | Indicates if the color of the background change between phases. The order is: - Blue (default) - Yellow - Orange - Purple After the last color, the order restarts. |
| Error_show_time | X | 0 to N | Indicates the time (seconds) during which the test shows feedback of the mistake when a participant selects an incorrect number. The feedback is that the rectangle turns red, the number flashes and the number buttons are disabled. |
| Correct_show_time | X | 0 to N | Indicates the time (seconds) during which the test shows feedback of the correct answer when a participant writes the completely correct number. The feedback is that the rectangle turns or stays green (depending on the previous number), the number flashes and the number buttons are disabled. |
| Time_feedback_deact_Opt | X | 1 to N | Indicates the time (seconds) during which the options are deactivated when a new trial begins so that the participant pays attention to the feedback regarding the performance in the previous trial (it appears at the left of the screen). |
| Time_decision_reminder | X | 1 to N | Indicates the number of seconds that must elapse during which the participant has not chosen an option so that the test presents the following reminder: <i>Please, make a choice to continue.</i> |
| Motivation_criteria | X | 0 to N | Indicates the number of clicks that the participant should make in the trial. If the participant does not make the minimum number of clicks, their scores cannot be considered. |
| Competence_criteria | X | 0 to N | Indicates the number of sums the participant has to solve so that his or her scores could be considered. |

Phase (j).

| Variable | Format | Values | Description |
|----------------------------|--------|---|--|
| Series_random | X | 1 = random 0 = not random | Indicates if the series of the phase are presented or not randomly |
| Color_button_trials_change | X | 0 = no color change 1 = color change in order 2 = color change random | Indicates if the color of the options change between series. The order is: - Green (default) - Light grey - Dark blue Turquoise After the last color, the order restarts |

Serie (s). We recommend that series have at least six trials each.

| Variable | Format | Values | Description |
|----------------------------------|--|--|--|
| Trials_series | X | 0 to N | Indicates the number of trials of the series |
| A_initial_Points | X | 0 to N | Indicates the number of points offered by option A at the start of the series. This value will be presented in the training trials (the forced trials) and in the first test trial. |
| B_initial_Points | X | 0 to N | Indicates the number of points offered by option B at the start of the series. This value will be presented in the training trials (the forced trials) and in the first test trial. |
| A_initial_Time_interval | X | 0 to N | Indicates the time interval of option A (for how many seconds will sums be presented) at the start of the series. This value will be presented in the training trials (the forced trials) and in the first test trial. |
| B initial_Time_interval | X | 0 to N | Indicates the time interval of option B (for how many seconds will sums be presented) at the start of the series. This value will be presented in the training trials (the forced trials) and in the first test trial. |
| Pair_forced_training_Trials | X | 0 to N | Indicates how many pairs of trials the participant is forced to choose one of the options at the beginning of the serie. (In one of the trials of the pair they will be forced to choose option B. In the other trial they will be forced to choose option A). 0= no forced trials 1= one pair, hence, two forced trials (The values of the options will be those indicated in A_Points, B_Points, A_Time_interval and B_Time_interval). These trials are training trials, not test trials. |
| Order_opt_forced_training_Trials | X | 0 = random 1 = Option A first 2 = Option B first | Indicates if the options presented in the forced trials are presented in a specific order or randomly. |
| A_points_trial | $i_1:X_1;$ $i_2:X_2;...$ $i_n:X_n$ | 1 to N | Indicates the number of points offered by Option A in the test trials. <i>i</i> indicates the number of the trial. X indicates the number of points that is offered Example: 3:6; 4:8 means that option B lasts 6 seconds in the 3 rd trial and 8 seconds in the 4 th trial. |
| A_time_trial | $i_1:X_1;$ $i_2:X_2;...$ $i_n:X_n$ | 1 to N | Indicates the number of seconds that elapses Option A in the test trials. <i>i</i> indicates the number of the trial. X indicates the number of seconds that it elapses Example: 3:6; 4:8 means that option B lasts 6 seconds in the |

| | | | |
|----------------|--|--------|---|
| | | | 3 rd trial and 8 seconds in the 4 th trial. |
| B_points_trial | $i_1:X_1;$ $i_2:X_2;...$ $i_n:X_n$ | 1 to N | Indicates the number of points offered by Option B in the test trials. <i>i</i> indicates the number of the trial. X indicates the number of points that is offered Example: 3:6; 4:8 means that option B offers 6 points in the 3 rd trial and 8 points in the 4 th trial. |
| B_time_trial | $i_1:X_1;$ $i_2:X_2;...$ $i_n:X_n$ | 1 to N | Indicates the number of seconds that elapses Option B in the test trials. <i>i</i> indicates the number of the trial. X indicates the number of seconds that it elapses. Example: 3:6; 4:8 means that option B lasts 6 seconds in the 3 rd trial and 8 seconds in the 4 th trial. |
| A_Bonus_trial | $i_1:X_1;$ $i_2:X_2;...$ $i_n:X_n$ | 1 to N | It indicates the points awarded for each extra sum (compared to the number of sums completed in the second trial of the series in which the participant chose option A) completed in option A. |
| B_Bonus_trial | $i_1:X_1;$ $i_2:X_2;...$ $i_n:X_n$ | 1 to N | It indicates the points awarded for each extra sum (compared to the number of sums completed in the second trial of the series in which the participant chose option B) completed in option B. |

Trial (i).

| Variable | Format | Values | Description |
|-----------------------------|--------|-----------------|---|
| Feedback_points_last_trial | X | 0: no 1: yes | Determines whether (1) or not (0) feedback on the points awarded in the last trial is presented in the left panel |
| Feedback_errors_last_trial | X | 0: no 1: yes | Determines whether (1) or not (0) feedback on the errors committed in the last trial is presented in the left panel |
| Feedback_seconds_last_trial | X | 0: no 1: yes | Determines whether (1) or not (0) feedback on the duration (seconds) of the last trial is presented in the left panel |
| Feedback_poins_accumulated | X | 0: no 1: yes | Determines whether (1) or not (0) feedback on the total points accumulated is presented in the left panel |

4. Data output

When executed, the data recorded by the test will be stored: general data, data recorded in each phase, serie, trial and global data. The data can be visualized as both plain text (.txt) and as a table (.xls).

General.

| Variable | Format | Values | Description |
|--------------------------------|--------|-----------------------------------|---|
| ID | X | 1 to N | Number that identifies the participant |
| Language | X | ES, EN | ISO 639-1 code that corresponds to the language in which the test was presented |
| Time_test | X | 1 to N | It indicates the duration (seconds) of the test. |
| N_phases | X | 1 to N | It informs about the number of phases of the test. |
| N_Series_phase_j | X | 1 to N | It informs about the number of series per phase of the test. |
| N_trials_phase_j_series_s | X | 1 to N | It informs about the number of trials of each series of each phase. |
| N_trials_test | X | 1 to N | It indicates the total number of trials of the test. |
| Motivation_criteria_fullfilled | X | 0= not fulfilled 1 = fulfilled | It indicates if the participant has met competence criteria in all the trials |

| | | | |
|--------------------------------|---|-----------------------------------|---|
| Competence_criteria_fullfilled | X | 0= not fulfilled 1 = fulfilled | It indicates if the participant has met competence criteria in all the trials |
|--------------------------------|---|-----------------------------------|---|

Phase (j).

| Variable | Format | Values | Description |
|-------------------------------------|--------|--------|---|
| Time_phase_j | X | 1 to N | It indicates the duration (seconds) of the phase |
| Latency_phase_j | X | 0 to N | Indicates the average latency of the phase (the mean time it takes for the participant to press an option) $\sum \text{latency_series_s} / J$ phases |
| Total_Sums_Completed_A_j | X | 0 to N | Indicates the total sums completed in the phase when working on option A |
| Total_Sums_Completed_B_j | X | 0 to N | Indicates the total number of sums completed in the phase when working on option B |
| Total_Sums_Completed_j | X | 0 to N | Indicates the total number of sums completed in the phase. |
| Total_Points_Sums_j | X | 0 to N | Indicates the total number of points accumulated in the phase. |
| Total_Errors_Sums_j | X | 0 to N | Indicates the total number of errors accumulated in the phase. |
| Opt_A_Total_j | X | 0 to N | Indicates the total number of trials the participant has chosen option A in this phase |
| Opt_B_Total_j | X | 0 to N | Indicates the total number of trials the participant has chosen option B in this phase |
| Efficacy_Total_Reinforcement_rate_j | X | 0 to N | \sum points of the options chosen / duration in seconds of the options chosen in this phase |
| Self_Control_j | X | 0 to N | Indicates the average self-control of the phase $\sum \text{Self_Control_s} / S$ series |

Series.

| Variable | Description |
|----------|--|
| Seq_i | Registers the actions (licks) of the user in the following format: a;b;c;d;e;f;g - 'a' is the moment (time in milliseconds) in which the participant clicks, - 'b' is the type of click (option A or B) - 'c' is the number of points achieved with that election - 'd' is the number of points that offered that option without the bonus - 'e' is the time needed to complete the chosen option - 'f' is the number of points that offered the opposite option - 'g' is the time needed to complete the opposite option |

| Variable | Format | Values | Description |
|--------------------------|--------|--------|---|
| Time_series_s | X | 1 to N | It indicates the duration (seconds) of the series |
| Latency_series_s | X | 0 to N | Indicates the average latency of the series (the mean time it takes for the participant to press an option) $\sum \text{latency_trial_i} / i$ trials |
| Total_Points_Sums_s | X | 0 to N | Indicates the total number of points accumulated in the series. |
| Total_Errors_Sums_s | X | 0 to N | Indicates the total number of errors accumulated in the series. |
| Total_Sums_Completed_A_s | X | 0 to N | Indicates the total sums completed in the series when working on option A |
| Total_Sums_Completed_B_s | X | 0 to N | Indicates the total number of sums completed in the series when working on option B |
| Total_Sums_Completed_s | X | 0 to N | Indicates the total number of sums completed in the series. |

| | | | |
|-------------------------------------|---|--------|--|
| Opt_A_Total_s | X | 0 to N | Indicates the total number of trials the participant has chosen option A in this serie. |
| Opt_B_Total_s | X | 0 to N | Indicates the total number of trials the participant has chosen option B in this serie. |
| A_Reinforcement_rate_s | X | 0 to N | \sum points of the options A chosen / duration in seconds of the options A chosen |
| B_Reinforcement_rate_s | X | 0 to N | \sum points of the options B chosen / duration in seconds of the options B chosen |
| Efficacy_Total_Reinforcement_rate_s | X | 0 to N | \sum points of the options chosen / duration in seconds of the options chosen |
| Self_Control_s | X | 0 to N | $10^* [1 + (\text{Points}_B_Reinforcement_rate_s) - (\text{Points}_A_Reinforcement_rate_s)]$ in the last 3 elections of the series |

Trial (i).

| Variable | Format | Values | Description |
|-------------------------------|--------|---------------------------------|--|
| Time_trial_i | X | -1, 1 to N -1: no time limit | It indicates how much time (seconds) the trial has lasted |
| Latency_trial_i | X | 0 to N | Indicates the latency (the time it takes for the participant to press an option) |
| Total_Sums_Completed_A_i | X | 0 to N | Indicates the total sums completed in the when working on option A |
| Total_Sums_Completed_B_i | X | 0 to N | Indicates the total sums completed in the trial when working on option B |
| Total_Sums_Completed_i | X | 0 to N | Indicates the total number of sums completed in the trial |
| Total_Points_Sums_i | X | 0 to N | Indicates the total points accumulated in the trial. |
| Total_Errors_Sums_i | X | 0 to N | Indicates the total number of errors accumulated in the trial. |
| Points_A_Reinforcement_rate_i | X | 0 to N | If the participant has pressed Option A: Points of the option A chosen/ seconds of the option A chosen If the participant has not pressed Option A: 0 |
| Points_B_Reinforcement_rate_i | X | 0 to N | If the participant has pressed Option B: Points of the option B chosen / seconds of the option B chosen If the participant has not pressed Option B: 0 |

Global data.

| Variable | Format | Values | Description |
|--------------------------|--------|--------|--|
| Latency_g | X | 0 to N | Indicates the average latency of the test (the mean time it takes for the participant to press an option) \sum latency_phase_j / j phases |
| Total_Sums_Completed_A_g | X | 0 to N | Indicates the total sums completed in the test when working on option A |
| Total_Sums_Completed_B_g | X | 0 to N | Indicates the total sums completed in the test when working on option B |
| Total_Sums_Completed_g | X | 0 to N | Indicates the total number of sums completed in the test |
| Total_Points_Sums_g | X | 0 to N | Indicates the total points accumulated in the test. |
| Total_Errors_Sums_g | X | 0 to N | Indicates the total number of errors accumulated in the test. |

| | | | |
|-------------------------------------|---|--------|---|
| Total_Opt_A_g | X | 0 to N | Indicates the total number of trials the participant has chosen option A in the test. |
| Total_Opt_B_g | X | 0 to N | Indicates the total number of trials the participant has chosen option B in the test. |
| Efficacy_Total_Reinforcement_rate_g | X | 0 to N | \sum points of the options chosen / duration in seconds of the options chosen in the test |
| Total_Self_control_g | X | 0 to N | Indicates the average self-control of the test \sum Self_Control_j / j phases |

5. Standard instructions and standard start screen (instr.json)

INSTRUCTIONS SCREEN 1: Instructions

- In this test, you will have to solve some simple arithmetic operations.
- We will give you the option to **choose** what kind of operations you want to solve.
- For solving the arithmetic operations, you will get points. Your goal is **to get as many points as possible**.

INSTRUCTIONS SCREEN 2: Instructions

Remember! Your *goal* is to *obtain the maximum number of points*.

START SCREEN: Are you ready?

If you are ready, press the start button and the test will begin.

If you prefer to review the instructions, click on the arrow below

6. Standard configuration (config.json)

[General]

Language = EN

Time_test = 2100

Feedback_Time_test = 1

Number_phases = 2

Number_series = 2

Number_series_phase_1 = 2

Number_series_phase_2 = 2

Number_trials_phase_1_serie_1 = 6

Number_trials_phase_1_serie_2 = 6

Number_trials_phase_2_serie_1 = 6

Number_trials_phase_2_serie_2 = 6

Color_background_phases_change = 1

Error_show_time = 2

Correct_show_time = 2

Time_feedback_deact_Opt = 3

Time_decision_reminder = 10

Motivation_criteria = 3

Competence_criteria = 1

[Phase 1]

Series_random = 0

Color_button_trials_change = 1

[Series 1]

Trials_serie = 6
A_initial_Points = 8
B_initial_Points = 30
A_initial_Time_interval = 6
B_initial_Time_interval = 60
Pair_forced_training_Trials = 1
Order_opt_forced_training_Trials = 1
A_points_trial = 1:8; 2:8; 3:8; 4:8; 5:8; 6:8
A_time_trial = 1:6; 2:12; 3:18; 4:24; 5:28; 6:32
B_points_trial = 1:30; 2:30; 3:30; 4:30; 5:30; 6:30
B_time_trial = 1:60; 2:60; 3:60; 4:60; 5:60; 6:60
A_Bonus_trial = 0
B_Bonus_trial = 0

[Trial 1]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 2]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 3]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 4]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 5]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 6]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Series 2]

Trials_serie = 6
A_initial_Points = 4
B_initial_Points = 60
A_initial_Time_interval = 6
B_initial_Time_interval = 40
Pair_forced_training_Trials = 1
Order_opt_forced_training_Trials = 1
A_points_trial = 1:4; 2:4; 3:4; 4:4 ;5:4; 6:4
A_time_trial = 1:6; 2:6; 3:6; 4:6; 5:6; 6:6
B_points_trial = 1:60; 2:60; 3:60; 4:60; 5:60; 6:60
B_time_trial = 1:40; 2:52; 3:64; 4:76; 5:86; 6:96
A_Bonus_trial = 0
B_Bonus_trial = 0

[Trial 1]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 2]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 3]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 4]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 5]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 6]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Phase 2]

Series_random = 0

Color_button_trials_change = 1

[Series 1]

Trials_serie = 6

A_initial_Points = 40

B_initial_Points = 150

A_initial_Time_interval = 6

B_initial_Time_interval = 60

Pair_forced_training_Trials = 1

Order_opt_forced_training_Trials = 1

A_points_trial = 1:40; 2:40; 3:40; 4:40; 5:40; 6:40

A_time_trial = 1:6; 2:10; 3:14; 4:18; 5:21; 6:24

B_points_trial = 1: 150; 2:150; 3:150; 4:150; 5:150; 6:150

B_time_trial = 1:60; 2:60; 3:60; 4:60; 5:60; 6:60

A_Bonus_trial = 2

B_Bonus_trial = 8

[Trial 1]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1

Feedback_seconds_last_trial = 1

Feedback_poins_accumulated = 1

[Trial 2]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1

Feedback_seconds_last_trial = 1

Feedback_poins_accumulated = 1

[Trial 3]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1

Feedback_seconds_last_trial = 1

Feedback_poins_accumulated = 1

[Trial 4]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1

Feedback_seconds_last_trial = 1

Feedback_poins_accumulated = 1

[Trial 5]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1

Feedback_seconds_last_trial = 1

Feedback_poins_accumulated = 1

[Trial 6]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Series 2]

Trials_serie = 6
A_initial_Points = 20
B_initial_Points = 300
A_initial_Time_interval = 6
B_initial_Time_interval = 40
Pair_forced_training_Trials = 1
Order_opt_forced_training_Trials = 1
A_points_trial = 1:20; 2:20; 3:20; 4:20; 5:20; 6:20
A_time_trial = 1:6; 2:6; 3:6; 4:6; 5:6; 6:6
B_points_trial = 1: 300; 2:300; 3:300; 4:300; 5:300; 6:300
B_time_trial = 1:40; 2:52; 3:64; 4:76; 5:86; 6:96
A_Bonus_trial = 1
B_Bonus_trial = 3

[Trial 1]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 2]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 3]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 4]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 5]

Feedback_points_last_trial = 1
Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

[Trial 6]

Feedback_points_last_trial = 1

Feedback_errors_last_trial = 1
Feedback_seconds_last_trial = 1
Feedback_poins_accumulated = 1

References

- Ashe, M. L., & Wilson, S. J. (2020). *A Brief Review of Choice Bundling: A Strategy to Reduce Delay Discounting and Bolster Self-Control*. *Addictive Behaviors Reports*, 100262. <https://doi.org/10.1016/j.abrep.2020.100262>
- Brigham, T.A. (1982). Self-management: A radical behavioral perspective. In P. Karoly & E.H. Kanfer (Eds.), *Self-management and behavior change: From theory to practice* (pp. 32-59). Pergamon.
- Cattell, R. B., & Warburton, F. W. (1967). *Objective Personality and Motivation Tests: A theoretical introduction and practical compendium*. University of Illinois Press.
- Rachlin, H. (1974). Self-control. *Behaviorism*, 2, 219-233.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control 1. *Journal of the experimental analysis of behavior*, 17(1), 15-22. <https://doi.org/10.1901/jeab.1972.17-15>
- Santacreu, J. & Hernández, J.M. (2017). T-Data (Tests). En V. Zeigler-Hill y T.K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp.1-4). Springer.