Current Research in Behavioral Sciences 1 (2020) 100005

Contents lists available at ScienceDirect

Current Research in Behavioral Sciences

journal homepage: www.elsevier.com/locate/crbeha

Researching unexplained phenomena II: new evidences for anomalous experiences supported by the Multivariable Multiaxial Suggestibility Inventory-2 (MMSI-2)

Álex Escolà-Gascón

Faculty of Psychology, Education and Sport Sciences, Blanquerna, Ramon Llull University, Barcelona, Spain

ARTICLE INFO

Keywords:
Parapsychology
Paranormal beliefs
Hallucinations
Anomalous experiences
Unexplained phenomena
Lie detection

ABSTRACT

Anomalous phenomena are human experiences that are characterized by challenging the foundations of current scientific ontology (i.e., psi phenomena). The problem lies in the fact that some studies have obtained significant results that support the existential validity of psi phenomena. This fact calls into question the role of psychology -and specifically that of psychological assessment- in scientifically justifying and objectively evaluating this type of behavior. This work examines the construct validity and reliability of the *Multivariable Multiaxial Suggestibility Inventory-2* (MMSI-2), a psychometric test that measures both anomalous phenomena and the main psychological predictive variables that could generate them. The study included 804 participants without psychiatric history. The participants were evenly distributed into two groups: participants who *believe* in the existence of the paranormal and participants who are *non-believers. Confirmatory factor analysis* was applied, *factorial invariance* between both groups was examined, and *Cronbach's alpha* and *Omega* reliability coefficients were calculated. The results allowed accepting the 'strong factorial invariance' for the internal structure of the MMSI-2. In parallel, *latent means* analysis indicated that believers had higher scores than non-believers in the 4 latent variables of the test. *Regression models* indicated that the *Clinical Personality Tendencies* (CPT), *Incoherent Manipulations* (IMA) and *Altered States of Consciousness* (ASC) scales predicted 51.2% of anomalous phenomena. It is concluded that the MMSI-2, with its 174 items and 20 scales, is a valid and reliable psychometric instrument. This research is a continuation of the Escolà-Gascón (2020) report, in which the first psychometric properties of the MMSI-2 were published.

1. Introduction

Certain types of behaviors that are scientifically difficult to explain are called *anomalous phenomena* (e.g., French and Stone, 2014), although they do not have to be inexplicable (see Lange et al., 2019). Research into these phenomena is complex because they challenge or might appear to contradict current scientific ontology (e.g., Parkinson, 2019). These phenomena can be very diverse and vary according to each scientific discipline (e.g., Bobrow, 1983, 2003). This report focuses on the psychometric study of *'psi' phenomena* and of anomalous experiences associated with parapsychology. The term 'psi' phenomena serves to classify the investigation of three objects of study (e.g., Irwin and Watt, 2007; Jinks, 2019): (1) *anomalous mind-to-mind communication* (also informally called "telepathy"); (2) *anomalous anticipation of information* (called "precognition"); and (3) *anomalous mind-matter interaction* (informally known as "psychokinesis") (see also Eysenck and Sargent, 1982). In some cases, other phenomena related to parapsychology are also included, such as *mediumship* or *out-of-body experiences* (hereinafter OBEs), which makes the classification of the 3 previous categories vary according to the criterion applied by professional researchers (e.g., Beischel and Zingrone, 2015). From psychiatry and clinical psychology, these behaviors are justified as hallucinatory symptoms (e.g., Kelly et al., 2020), perceptual alterations or bias (e.g., Wright et al., 2020) and as belief systems that allow the attribution of "paranormal" meanings to the daily experiences that each subject experiences (e.g., Irwin, 2009, 2003; Irwin et al., 2013; Jinks, 2019). For this reason, in place of informal terminology, the use of the expression "anomalous phenomena" or "anomalous behavior" is accepted. On the one hand, they are behaviors whose clinical or psychopathological value is unclear (e.g., David, 2010; Nordgaard et al., 2019) and, on the other hand, they also assume the hypothesis that some unknown psychological mechanism intervenes in the development of these behaviors (e.g., Utts, 2018). This hypothesis is called the *'psi' hypothesis* and differs from the paranormal model in that it does not assume the existence of supernatural forces or realities (e.g., Mayer, 2017). However, many researchers confused this hypothesis and understood it as one more expression of beliefs in the paranormal (see Carter, 2012).

In any case, anomalous phenomena are observable behaviors in psychiatric and psychological evaluations (e.g., Parker, 2006; Shapiro et al., 2019). Because of this, psychology and psychiatry play a role in how to

E-mail address: alexeg@blanquerna.url.edu

https://doi.org/10.1016/j.crbeha.2020.100005
Received 5 October 2020; Received in revised form 27 October 2020; Accepted 10 November 2020

evaluate, quantify and identify the criteria that should be used to scientifically explain this type of behavioral anomaly (e.g., Lawrence, 2016). There is sufficient evidence that discredits or questions the scientific validity of 'psi' phenomena (see O'Keeffe and Wiseman, 2005; Reber and Alcock, 2019; Wagenmakers et al., 2011). However, this type of phenomenon is not incompatible with the scientific method, and numerous studies present significant results in favor of the 'psi' hypothesis (e.g., Beischel et al., 2015; Bem, 2011; Bem et al., 2016; Honorton, 1985; Kelly and Arcangel, 2011; Maher, 1999, 2000, 2015; Maher and Hansen, 1992, 1995; Mossbridge et al., 2012; Robertson and Roy, 2001, 2004; Roy and Robertson, 2001; Schwartz and Russek, 2001). Thus, it is no longer a debate exclusive to the "philosophy of science." The fact that there is scientific research with results that support the validity of the 'psi' hypothesis makes the scientific discussion of these objects of study also methodological (e.g., Jinks, 2019). According to Tressoldi and Utts (2015), this has three main implications: (1) the systems used in scientific research to measure and quantify the 'psi' phenomena must be examined; (2) the methodological designs and statistical analyses used should be reviewed; and (3), the procedures and results of the investigations should be replicated. Therefore, all this requires the application of the scientific method in the evaluation and examination of anomalous phenomena. It should be noted that the method used can be applied at different levels and ways (e.g., Bunge, 2013; Wright and Hallquist, 2020). One of these ways or levels is the psychometric approach, especially using self-reporting techniques (e.g., Abad et al., 2015).

One of the problems is how to evaluate the behavior of an individual in an objective and scientific manner (e.g., Groth-Marnat, 2009). Although the method of direct and systematic observation is applicable in clinical psychology, the technique most commonly used in evaluation is indirect observation through structured self-reporting tests (e.g., Miller and Lovler, 2020). These self-reporting tests have several classifications (see Weiner and Greene, 2017), but the two most commonly used are *structured interviews* and *self-report questionnaires*. The self-report questionnaires allow the psychometric profile to be traced with the scores associated with the psychological variables that could explain a certain type of behavior/discomfort (e.g., Kline, 2013). There are multiple questionnaires or psychometric instruments that serve to quantitatively measure anomalous phenomena (e.g., Bell et al., 2006; Mason and Claridge, 2006; Stefanis et al., 2002).

In the context of anomalous experiences, self-report questionnaires have been used for three main purposes: (1) as measures to quantify **hallucinatory behaviors** and to try to discriminate the most intense (pathological) symptomatology from the most attenuated symptomatology (not pathological) (e.g., Johns and van Os, 2001; Shapiro et al., 2019; van Os et al., 2009). This idea is based on the *psychosis continuum* model and is not only applied with hallucinatory symptoms but also includes other traits attributed to psychotic symptoms, which are summarized in the so-called *negative symptoms* (see Fekih-Romdhane et al., 2020). The purpose of evaluating hallucinations refers to identifying the *psychotic phenotype* that would allow estimating and preventing future psychotic crises (e.g., Oliver et al., 2019). (2) They have also been used to quantify **perceptual distortions or alterations** and **cognitive biases** (e.g., Barberia et al., 2013; Barberia et al., 2018; Groome et al., 2019; Khun et al., 2016). This use comes from the *semiotic model of perception* and from cognitive psychology, which determine a very clear difference with respect to hallucinations: in perceptual distortions, a provocative sensory object exists, but it is perceived in an altered way (e.g., Belloch et al., 1995). In contrast, the hallucination develops without provocative sensory stimuli, and it is the patient who infers new unreal perceived content (e.g., Upthegrove et al., 2015). The objective of these questionnaires is related to basic science applied to the psychology of perception and the exploration of illusory symptoms (which are not hallucinations) present in both psychotic symptoms (e.g., Chapman et al., 1978) as in other non-psychotic clinical conditions, such as *eating disorders* (e.g., Sirvent et al., 2019). (3) Another use is found in the **measurement of belief systems and mental representations** associated with the paranormal (e.g., Font, 2016; Heotis, 2019; Irwin, 1993, 2000, 2003, 2009; Jaspers, 1993). This is based on the *phenomenological model* and on the part of cognitive psychology that studies mental processes related to consciousness (e.g., French and Stone, 2014). This model postulates that certain scientifically impossible phenomena—for example, *hearing the thoughts of another person* (which would be an experience of mind-to-mind interaction)—are not explained by hallucinatory or perceptual errors (e.g., Irwin, 2009). According to Irwin et al. (2013) this type of experience is a subjective cognitive attribution that generates a magical interpretation of an ordinary and extraordinary situation. The concept of "extraordinary" describes uncommon but scientifically possible situations. An example is the phenomena of random coincidences (e.g., *having the feeling that something bad has happened to someone and that later is true*). It is possible that the presentiment is explained both by the hypothetical-deductive rational processes developed by the subject, as well as by processes of a more emotional or intuitive nature (e.g., Jinks, 2019; Parkinson, 2019). In any case, it is the individual who mentally represents this phenomenon under an attribution that can be inclusive within the framework of science or exclusive to the scientific world (which in this case would be the "paranormal" attribution) (e.g., Drinkwater et al., 2017). The importance of this model lies in the type of representation that the subject produces and not in whether the phenomenon occurred as described by the patient (e.g., Cameron, 2016; Font, 2016).

For the 'psi' phenomena, the following detail should be clarified: None of the three previous points accepts the ontological and scientific validity of 'psi' phenomena (e.g., Reber and Alcock, 2019). In reality, in the psychological evaluation, it is not necessary to check whether the anomalous phenomenon described by the patient has a direct empirical reference, but its irrational and divergent content with scientific discourse should not necessarily verify its hallucinatory condition (e.g., Bobrow, 2003). This is a clear example of the *Aristotelian fallacy of verifying a consequence from an uncertain cause* (e.g., Pardo and Román, 2013). Assuming that it is a "perceptual alteration" (consequence) because "the discourse seems incompatible with the rational principles of science" (antecedent) is equivalent to the following fallacious statement: "the ground is wet (consequence) because it has rained (antecedent)". The antecedent is uncertain because the ground may have gotten wet in many other ways. Confirmation that it had rained would also not exclude other possibilities; for example, someone might have previously been washing with water. The same happens with the 'psi' phenomena evaluated in psychiatric practice: it can be accepted that there are "perceptual alterations" (consequence), but not because their content is "scientifically impossible" (antecedent). Other possible hypothetical precedents may exist within the scientific framework, such as that the patient simulates or undertakes fraud (e.g., Leonard and Williams, 2019). Another possibility would be cognitive biases such as the *Barnum effect* (e.g., Shermer, 2011) or the systems of meaning themselves (see Irwin, 2009).

All this indicates that when faced with an anomalous experience, especially a supposed 'psi' phenomenon, the causes should not be judged from the "diagnostic impression" (e.g., Parker, 2006). Even in a clinical examination, the hypothetical causal antecedents must be scientifically contrasted using the tools provided for this purpose (e.g., Groth-Marnat, 2009). It is at this point that several problems arise.

First, in the scientific literature, there are no psychometric tools that assess anomalous phenomena, including the three applications described in the previous paragraphs (e.g., Houran et al., 2019). Psychometric scales can be found that separately and independently examine anomalous phenomena such as hallucinations, distortions or perceptual illusions and belief systems (e.g., Wahbeh et al., 2019). The main drawback is that they are analyzed as if they were independent psychiatric or psychological models, when in reality they are correlated with each other (e.g., French and Stone, 2014). Second, although these specific scales present satisfactory statistical validation, none of them explore other possible psychological antecedent variables that can correlate with the anomalous experiences experienced. Knowing and analyzing these concomitant variables is something that could also contribute to psychological intervention, since they would give clues to the health professional about what dimensions should be modified to achieve a therapeutic change (e.g., Harary, 2006). Third, while it is true that clinical questionnaires are available to examine standardized psychiatric traits and symptoms (e.g., Butcher et al., 2019), it is also true that the items from these instruments were originally designed and validated with a medical-pathological population (e.g., Morey, 2011). This is another drawback because although the items and their scores are measured with non-clinical control groups, their content will remain pathological because it will not vary qualitatively (e.g., Butcher et al., 1995; Fernández-Ballesteros, 2011). This impairs the professional evaluation of anomalous phenomena in the general population because according to this logic, the experiences would be assumed to be exclusively pathological hallucinations (e.g., Parker, 2006). As Pasricha (2011) warns, classic clinical inventories are tools that in this study object provide biased information that does not help professional researchers make effective decisions. New tools are needed to explore the main hypothetical data and help the researcher or health professional make more accurate and objective clinical decisions (e.g., Waugh et al., 2017; Wright and Hallquist, 2020).

This report examines the confirmatory validity and reliability indices of the *Multivariable Multiaxial Suggestibility Inventory-2* (MMSI-2). Specifically, the factorial invariance of the MMSI-2 is analyzed in a group of subjects who believe in the paranormal and another group formed by non-believers. The purpose is to know if the believing subjects tend to score higher in the MMSI-2 than the non-believers, and should that be the case, if the reasons for these higher scores are related to the act of believing in the paranormal or to the MMSI-2. The MMSI is a psychometric instrument that quantifies anomalous phenomena by integrating them and relating them with 12 other psychological variables: *Inconsistencies* (K), *Lies* (L), *Fraud* (F), *Simulation* (Si), *Neurasthenia* (Nt), *Substance Use* (Cs), *Suggestibility* (Su), *Thrill-Seeking* (Be), *Histrionism* (Hi), *Schizotypy* (Ez), *Paranoia* (Pa) and *Narcissism* (Na). The test groups the anomalous phenomena into four dimensions: *Anomalous Visual/Auditory Phenomena* (Pva), *Anomalous Tactile Phenomena* (Pt), *Anomalous Olfactory Phenomena* (Po) and *Anomalous Cenesthetic Phenomena* (Pc). The objective of this report lies in contrasting the theoretical structure of MMSI-2, which is based on *exploratory factor analysis* (EFA) applied previously by Escolà-Gascón (2020) in a sample of more than 3,000 subjects. According to the results of the first validation phase, the 16 dimensions were distributed into 4 latent variables: *Clinical Personality Tendencies* (CPT), *Anomalous Perceived Phenomena* (APP), *Incoherent Manipulations* (IMA) and *Altered States of Consciousness* (ASC). Therefore, we intend to analyze the statistical relationship between the 4 latent variables and whether the CPT, IMA and ASC factors are correlated with perceptual alterations (APP).

2. Methods

2.1. Participants

The subjects who were part of this study came from the Spanish general population, specifically from Madrid (50%) and Barcelona (50%) (N_{total}=804). The participants signed an informed consent form, voluntarily collaborated with the study without receiving any financial compensation and reported no psychiatric history. The latter was the main inclusion criterion, since if the participants had a psychopathological history, the probability of suffering from a mental disorder would be higher and, therefore, also the probability of belonging to a medical population group. The sampling was not probabilistic.

The 804 subjects were classified according to two groups: (1) *believers in the existence of the paranormal* (N = 402) and (2) *nonbelievers in the existence of the paranormal* (N = 402). This classification was made from self-reported data that participants declared about their belief system. Each participant was asked the following question: *Do you believe that paranormal phenomena exist?* The subjects who answered 'yes' were part of the 'believing group' and those who answered 'no' formed the 'non-believers'. In each group, the number of subjects associated with the variables *sex* (men and women), *educational level* (classified according to the *National Statistics Institute of Spain*) and city of residence (Madrid or Barcelona) were equal to 50% except for 33.3% in the case of 'educational level'. This descriptive information is specified in Table 1.

For the analysis of the *age* variable, the *means* (M) and the *standard deviations* (SD) were calculated for each specific group of subjects. Again, Table 1 summarizes the information of these statistics. The age of the participants did not show significant differences in their means between the group of believers and non-believers. This ensured that the means associated with age were similar in each sample and that they their dispersion was homogeneous.

2.2. Procedure

This study is based on multivariate and *ex post facto* research designs, mainly using *structural equation models* and statistical analysis of *invariance*. The preparation of the research, data collection, statistical analysis and the report presented here are part of a university project that aims to measure statistics and psychological prediction of anomalous phenomena. This project began in 2013 and triggered the development of the *Multivariable Multiaxial Suggestibility Inventory-2* or MMSI-2, a new psychometric test that evaluates anomalous phenomena and the main psychological indicators that could justify them scientifically (see Escolà-Gascón and Gallifa, 2020). The construction of the MMSI-2 and part of its statistical justification are described in Escolà-Gascón (2020). This fact is relevant because it supports the beginnings of the research presented here.

To summarize, the wording of all the items of the MMSI was based on the scientific literature that related abnormal experiences with certain psychological attributes (e.g., Irwin, 2009). No initial or previous theoretical model was used. Decisions about which items should be written and what possible scales could be developed in the MMSI were determined by the empirical evidence published in the current scientific literature. The main psychological variables significantly correlated with anomalous phenomena were identified, and the MMSI items were written based on their characteristics or properties. Once the first inventory of items was developed (called the MMSI), each item was reviewed and analyzed using a content validity process. Of the 223 initial items, 49 were eliminated. The remaining 174 items were applied to a large sample of subjects belonging to the Spanish general population (N = 3,224). By not having a prior theoretical model, it was decided to contrast its empirical value and internal structure by applying various *exploratory factor analyses* (EFAs). The factorial solutions determined which groups of items would form the scales of the test. Then, using the direct scores of these scales, an initial scale and standardization of the MMSI with differentiated normative groups according to the variable 'sex' was performed. These first analyses concluded with the MMSI-2 version. The present research is based on these analyses and examines the construct validity of the MMSI-2 according to the empirical-statistical scales constructed from the initial EFAs (see Escolà-Gascón, 2020). Fig. 1 shows a diagram that summarizes the dimensions of the MMSI-2, its macro-factors and scales.

The hypothetical model aims to correlate the anomalous phenomena grouped in the latent variable APP with the other variables. Fig. 1 only represents the hypothesis of the model to contrast based on the initial EFAs.

After defining the structural model and the measurement model, in July 2018, the methodological preparation of the sampling and application materials began, and the collaborating professionals who would be responsible for the collection of the sample were contacted (see the acknowledgments section). At this point, the informed consents were also drafted, it was ensured that the recorded data of the participants was anonymous and the survey was designed to identify the previous so-

Table 1
Subject recounts and sample settings.

	Groups	Initial sample (N= 946)	Removal process (Removed= 142)	Final sample (N=804)
Sex				
M	Believers	217	16	201
	Non-believers	266	65	201
W	Believers	217	16	201
	Non-believers	246	45	201
Education level				
A	Believers	144 (M= 50; W= 94)	10 (M= 4; W= 6)	134 (M= 46; W= 88)
	Non-believers	163 (M= 58; W= 105)	29 (M= 23; W= 6)	134 (M= 35; W= 99)
B	Believers	138 (M= 67; W= 71)	4 (M= 4; W= 0)	134 (M= 63; W= 71)
	Non-believers	175 (M= 66; W= 109)	41 (M= 23; W= 18)	134 (M= 43; W= 91)
C	Believers	152 (M= 100; W= 52)	18 (M= 8; W= 10)	134 (M= 92; W= 42)
	Non-believers	174 (M= 142; W= 32)	40 (M= 19; W= 21)	134 (M= 123; W= 11)
Spanish Cities				
BCN	Believers	219 (M= 148; W= 71)	18 (M= 11; W= 7)	201 (M= 137; W= 64)
	Non-believers	262 (M= 185; W= 77)	61 (M= 34; W= 27)	201 (M= 151; W= 50)
Mad.	Believers	215 (M= 69; W= 146)	14 (M= 5; W= 9)	201 (M= 64; W= 137)
	Non-believers	250 (M= 81; W= 169)	49 (M= 31; W= 18)	201 (M= 50; W= 151)
Age	Believers	Mean= 27.72; SD= 12.454	*t* test= -0.076; *df*= 802; *p*= 0.939 *U* test= 79,079.5; *p*= 0.599 *Z* test= -0.526; *p*= 0.599	
	Non-believers	Mean= 27.65; SD= 12.482		

Note: *A= Compulsory secondary education or basic vocational training; B= Baccalaureate or higher vocational training; C= university or higher education; M= Men; W= Women; BCN= Barcelona; Mad.= Madrid; df= Degrees of freedom; SD= Standard Deviation.*

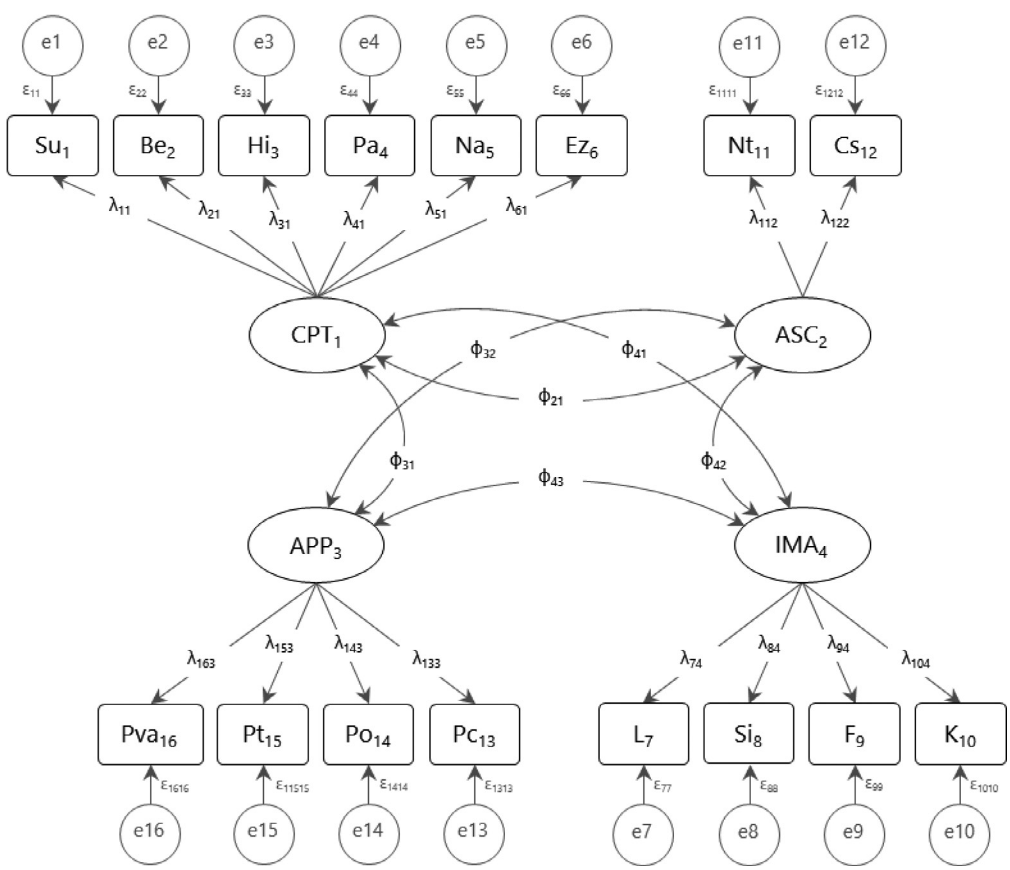

Fig. 1. Hypothetical model of the internal structure of the MMSI-2 using structural equations.

ciodemographic data. The direct scores obtained for each MMSI-2 scale were recorded in the data matrix. Although in the exploratory validation of the MMSI, psychometric scores were already made based on the T scores, the purpose of this research was statistical and was not intended to question or define individual interpretation criteria that require the study of the metric quality of the scales. Therefore, we chose to work statistically with the direct scores instead of using the T scores, which were still recently obtained and lacked subsequent statistical replications. In the same vein and following the example of Arribas (2011), the responses of the 174 items of the MMSI-2 were not recorded because each professional was provided with a system of correction templates that automatically allowed obtaining the direct scores for each factor. This was done to facilitate sample collection and computation of the data.

In September 2018, formal data collection began. Between January and March 2019, data entry began, and an initial sample sufficiently large (N = 359) was collected to carry out a prior statistical check. Statistical normality tests were performed, the heterogeneity of the data was analyzed, and the correlation matrix was examined. These first analyses supported the idea of applying a CFA and continuing with expanding the sample. It was at this point when it was decided to continue collecting data to apply CFAs in different types of samples. This converted the research design into a *multiple-group model*. This design offers the possibility of performing a statistical analysis of *configural invariance, weak factorial invariance (metric), strong factorial invariance (scalar)* and *strict factorial invariance (residual)* (see Brown, 2015).

In December 2019, the data collection phase ended. In total, 946 subjects participated (42% men and 58% women). During that month, the statistical control technique of *equalizing* proportions over recorded sociodemographic variables was applied (see Fleiss et al., 2003). This information is summarized in Table 1. First, the number of subjects was matched according to the variable *beliefs in the existence of the paranormal*. Two groups were established, one consisting of subjects who claimed to be believers in the existence of the paranormal (N = 476) (believing subjects) and another with individuals who did not believe in the paranormal (N = 476) (unbelieving subjects). Second, the number of subjects was also balanced according to the variables *sex* (men and women); *educational level* (compulsory secondary education or modules of basic professional training, baccalaureate studies or higher professional training and university or higher education); and *place of residence* (Barcelona or Madrid). Taking into account the objective of the 'equalizing technique', a total of 142 subjects were eliminated using 4 steps: (1) First, counts for each subgroup were performed. (2) The decision of how many subjects should be excluded to match the counts was determined by the *minimum observed frequency of subjects* in the subgroups classified in Table 1. The smallest observed frequency was that belonging to the variable 'educational level' for category B and for the believers sub-group, which had a recount of 138 subjects. (3) The third step was related to *outliers*. At this point, it was observed that there were 4 subjects with atypical direct scores on the scales of the L (Lies), Pva (Anomalous Visual/Auditory Phenomena) and K (Inconsistencies) tests. These scores were below the minimum value that can be obtained in the respective scales. Therefore, these 4 subjects were eliminated from believers sub-group B, reducing its number to 134. As it was the smallest frequency within the variable 'educational level', the value 134 was the corrected minimum value (MV) that all subjects should have distributed in that variable, since:

$$MV_x = \left(MO_{ij} - AV_{ij}\right)$$

$$MV_{EL} = \left(MO_{BB} - AV_{BB}\right)$$

$$= (138 - 4) = 134 \cdot 6 = 804$$

where

MV_x is the corrected minimum value of the *counts* of the *x* variable 'educational level' or *EL*;

MO_{ij} is the minimum observed frequency in the *i* group and in the *j* subgroup of the selected variable, which in this case is educational level *B* and the subgroup is that of believers (now also *B*); and

AV_{ij} is the number of cases with outliers observed in the *i* group and in the *j* subgroup of the selected variable, which in this case is educational level *B* and subgroup *B*.

The matching of these 6 subgroups to 134 subjects yields a total number of 804 subjects, which was also used to balance the groups of the other sociodemographic variables in parallel. (4) At this time, the fourth filter was applied simultaneously. Knowing that in the variables 'sex' and 'city of origin' the distinction was made between believing and non-believing subjects, in total, there should be 4 sub-groups in each of the two previous variables (see Table 1). If the final sample had 804 subjects in total, each of these 4 groups had 201 subjects. Thus:

$$O_{ij} - O'_{ij} = O_{1B} - O'_{1B} = 219 - 201 = 18$$

$$O_{1NB} - O'_{1NB} = 262 - 201 = 61$$

$$O_{2B} - O'_{2B} = 215 - 201 = 14$$

$$O_{2NB} - O'_{2NB} = 250 - 201 = 49$$

where

O_{1B} is the observed frequency belonging to BCN (1) and the subgroup of believers (B)

O_{1NB} is the observed frequency belonging to BCN (1) and the subgroup of non-believers (NB).

O_{2B} is the observed frequency belonging to Madrid (2) and sub-group B.

O_{2NB} is the observed frequency belonging to Madrid (2) and subgroup NB.

O'_{ij} are the relative frequencies of the assigned number of subjects (201).

(5) Finally, for each of the previous results (also applied to the variable 'educational level'), random quantities of men and women were eliminated. In total, 81 women and 61 men were eliminated. It was not possible to match the variable sex in the eliminations because the original counts did not allow it. It is important to highlight that eliminating the men and women from the cases was done randomly to eliminate the previously selected variables of city of residence and 'educational level'. The selection of these cases was not completely random since the presence of outliers (as indicated in step 3) was also taken into account. Thus, of the 18 subjects selected to be eliminated in the case of O_{1B}, 2 of them contained outliers. The remaining 16 subjects were randomly eliminated. This logic was applied in the rest of the eliminations. In total, the initial sample had 23 cases with outliers that were eliminated. In the remaining cases (of the 119 remaining subjects), the elimination of men and women was random. In any case, the final sample (N = 804) did not have cases with atypical scores. The variable 'age' was analyzed by comparison of means (*t-test*) and ranges (*Mann-Whitney U test*).

After the statistical cleansing of the original data matrix, data analysis and application of the structural equation models were developed, and the present research report was drafted.

2.3. Instruments

We used the *Multivariable Multiaxial Suggestibility Inventory-2* (MMSI-2), composed of 174 polytomous items distributed in the following scales: *Inconsistencies* (K), *Lies* (L), *Fraud* (F), *Simulation* (Si), *Neurasthenia* (Nt), *Substance Use* (Cs), *Suggestibility* (Su), *Thrill-Seeking* (Be), *Histrionism* (Hi), *Schizotypy* (Ez), *Paranoia* (Pa), *Narcissism* (Na), *Anomalous Visual/Auditory Phenomena* (Pva), *Anomalous Tactile Phenomena* (Pt), *Anomalous Olfactory Phenomena* (Po) and *Anomalous Cenesthetic Phenomena* (Pc). It also has scales elaborated empirically from different second-order factor analyses: *Clinical Personality Tendencies* (CPT), *Anomalous Perceived Phenomena* (APP), *Incoherent Manipulations* (IMA) and *Altered States of Consciousness* (ASC).

The participant must indicate up to what point he/she considers the contents of each item to be true using a 5-point *Likert* scale, which ranges between 1 (which means *completely disagree*) and 5 (which means *completely agree*). All questions must be answered. If the subject leaves items unanswered, this could lead to outliers in the dimension scores of the test (below the minimum direct score of each scale). If this occurs and cannot be resolved, the outliers should be invalidated or the entire profile should be excluded. In the Spanish version, the direct scores of each scale can be transformed to standardized scores (or T scores), which facilitate the individual analysis of scores and the preparation of psychological profiles.

The results of the test can be interpreted gradually; as the scores in each dimension increase, the greater the probability that the subject will present the respective attribute measured.

2.4. Data analysis

In relation to construct validity, *structural equation models* were used to contrast and analyze the theoretical structure of the MMSI-2 obtained empirically by statistical methods based on EFAs. More specifically, following the statistical recommendations of Brown (2015), the *confirmatory factor analysis* (CFA) technique was applied, and it was decided to develop a *cross-validation* psychometric design with the two samples of participants (group of believers and non-believers). We wanted to test whether the structural model of MMSI had acceptable and equivalent construct validity in both groups. When the variables are correctly measured and represented in the items of the inventory, the equality constraints imposed by the model (*configural invariance)*, the factor loadings (*weak factorial* (metric) *invariance*), the item intercepts (*strong factorial* (scalar) *invariance*) and residual variations (*strict factorial* (residual) *invariance*) should not impair the goodness of fit of the non-constrained model. According to Byrne (2014), one should test whether the changes that these constraints cause in the fit indices of the unconstrained model are significant. It is essential that the variations in the fit indices are not significant because in the contrary case, it would indicate the possibility that the contrasted models in each group would be different. This would imply questioning whether in both groups the same construct would actually be measured and, therefore, would also entail questioning the construct validity of the questionnaire. For the application of the *multiple-group* analysis based on 'strong factorial (scalar) invariance' and 'strict factorial (residual) invariance', the following steps were developed, specified by Brown (2015): (1) The hypothetical model of Fig. 1 was applied to each group separately using the CFA technique. The purpose was to ensure that the model in Fig. 1 fit satisfactorily in both groups. (2) Once the fit indices for each group of subjects were accepted, a CFA was again applied to the complete set including the two groups with and without invariance constraints. (3) Upon observing a minimal variation in the fit indices, we decided to contrast the null hypothesis of statistical significance. For this, the *Chi Square* statistic was used. Given that this statistic is highly sensitive to sample size, the criterion developed by Vandenberg and Lance (2000) was used, based on the cutoff point 0.01 in the comparative indices, specifically the *comparative fit index* (CFI). If the application of a determined invariance constraint caused this index to vary in an amount equal to or greater than 0.01, then the null hypothesis of equality would be rejected, and the result would suggest that the structural model is not the same in the groups. The equality between the empirical *covariance matrices* of the two groups was not analyzed because this statistical contrast can yield errors and results that would contradict the subsequent invariance analysis (e.g., Byrne, 2014; Byrne et al., 1989; Jaccard and Wan, 1996). As a preliminary exploration, it is worth noting that prior to the CFAs, EFAs were applied to ensure whether it was recommendable to proceed with the subsequent structural equations. All CFAs were based on the *maximum likelihood* method, and EFAs were calculated using the *unweighted least squares* method. Likewise, these analyses were carried out with the statistical program *SPSS.25* and its extension *AMOS*, specializing in structural equation models.

Reliability was analyzed based on the *Cronbach's alpha coefficient* (α), taking into account that the MMSI scales are quantitative interval variables. This statistic examines the *internal consistency* of the scores from the *variance-covariance matrix* between the scales of the test. However, the *Cronbach's alpha* becomes unstable when the matrix dimension is 2×2 and penalizes the consistency when the values of the variables are excessively heterogeneous (see Abad et al., 2015). As an alternative and complementary analysis, the calculation of *Omega's reliability coefficient* (ω) was applied. Although there are numerous estimates, all of them are based on the model factor loadings and on the *communality* of the items (see Trizano-Hermosilla and Alvarado, 2016). Mathematical expression (1) was chosen in this study, which is based on the contributions of McDonald (1999):

$$\omega_t = \frac{\left(\sum \lambda_j\right)^2}{\left[\left(\sum \lambda_j\right)^2 + \sum\left(1-\lambda_j^2\right)\right]} = \frac{\left(\sum \lambda_j\right)^2}{\left[\left(\sum \lambda_j\right)^2 + \left(\sum \psi\right)\right]} \quad (1)$$

where λ_j is the factor loading of item *j*,

λ_j^2 is the communality of item *j*, and
ψ is the unique variance.

This equation is incorporated into the statistical program JAMOVI (see The Jamovi Project, 2019), which is open access and was the one used for this calculation. All reliability coefficients were calculated for each group (believers and non-believers) and for the total sample.

3. Results

3.1. Initial exploratory factor analysis

Before the application of the respective CFAs, a descriptive and factorial exploratory examination of the working groups included in this research proceeded. The first point to know was whether the two groups (believers and non-believers) had similar scores and trends in the first-order scales of the MMSI. To do this, various (parametric and non-parametric) statistical hypothesis tests were calculated. All of them are specified in Tables 2 and 3. The effect size of each contrast was also examined by the *Cohen's d* (see Cohen, 1988). Significant results were obtained in the 16 scales/dimensions that suggested rejecting the null hypothesis of equality of means. Believers tended to systematically score higher than non-believers in the existence of the paranormal. The scales that obtained the largest effect sizes were dimensions L, Nt, Hi, Ez, Na and Pva. The Be scale yielded the smallest effects. Measures of central tendency and dispersion support these results. One unusual result can be observed in the *Mann-Whitney U test* of the L scale, whose critical value tends to infinity and therefore is truncated to 0. This also indicates that the difference between the means (more specifically between *interquartile* ranges) exceeds the standard deviation of the group of non-believers six-fold.

The EFA applied in each of the groups shows a similar factorial solution between the two types of samples. Unlike CFAs, in EFAs, it is advisable to apply the *unweighted least squares* extraction method, since it is the most conservative and allows the previous calculation of the 'communalities' necessary to deduct the subsequent factor loadings (e.g., Mulaik, 2018). As seen in Table 4, in each group, 4 latent factors were extracted with eigenvalues greater than 1. The EFA applied to the group of believers explained 87.75% of the total variance. In contrast, the group of non-believers yielded a total explained variance of 81.91%. In this same group, it seems that the K, L, F and Si scales also have a high saturation in the first factor (CPT). However, the loads of these dimensions remain highest in the third factor (IMA). These preliminary results favor the use of CFAs through structural equations and indicate that it is recommended to contrast the internal structure of the MMSI in both groups.

3.2. Multiple-group structural equation models

For each group, the CFA technique was applied based on the structural model of Fig. 1. Figs. 2 and 3 show the regression coefficients and standardized covariances between the respective variables of the model. As can be observed, the factorial coefficients are similar between both groups. Regarding the goodness of fit, both models presented indices with satisfactory values (see Table 5). According to Abad et al., (2015) and Kline (2013) the following adjustment indices thresholds were used: root mean square error of approximation (RMSEA, threshold= <0.05); comparative fit index (CFI, threshold= >0.95);

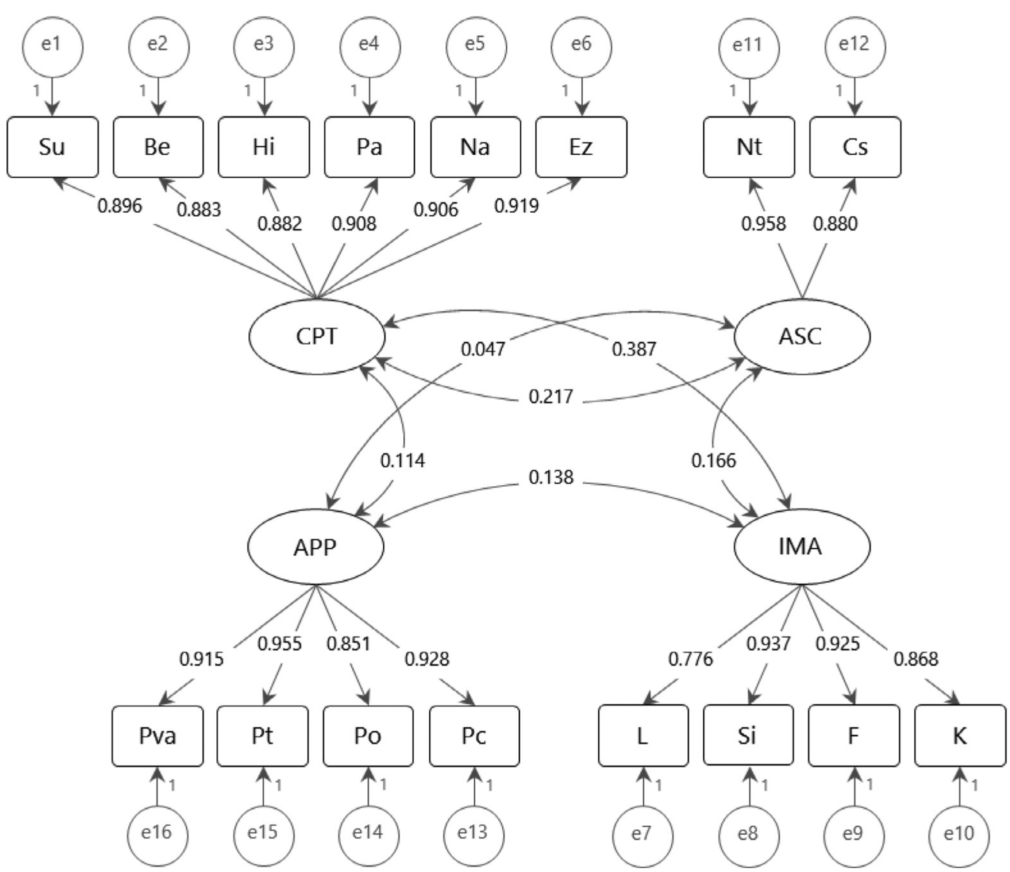

Fig. 2. Weighted regressions and standardized covariances between the variables of the MMSI-2 theoretical model (group of believers).

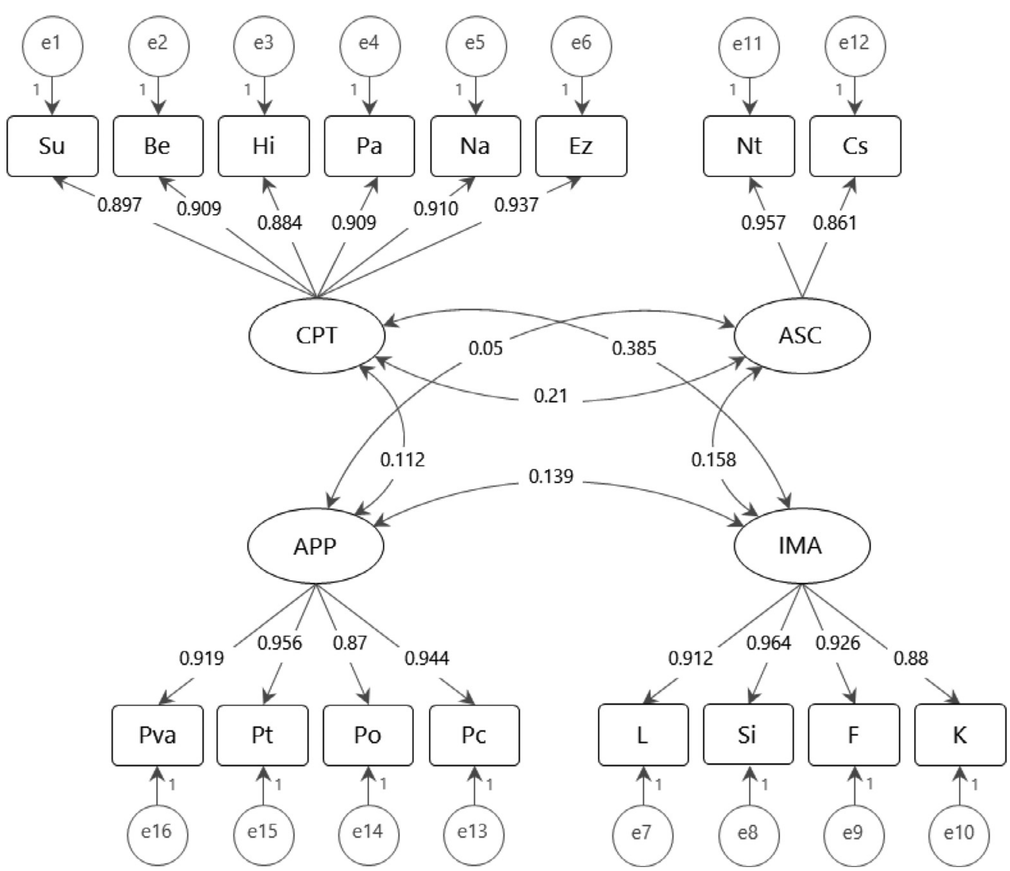

Fig. 3. Weighted regressions and standardized covariances between the variables of the MMSI-2 theoretical model (group of non-believers).

Table 2
Descriptive statistics of MMSI-2 scales.

Scales	Believers		Non-believers		Complete sample		*t test* *U test* *Z test*	*Cohen's d*
	M	SD	M	SD	M	SD		
K	21.28	2.628	17.32	2.595	19.30	3.279	−21.538 (*p*<0.001) 22,647 (*p*<0.001) −17.736	1.516**
L	55	3.347	33.86	2.97	44.43	11.039	−94.724 (*p*<0.001) ∞ ⇒ 0 (p<0.001) −24.573	6.681**
F	42.53	2.862	31.56	2.853	37.05	6.186	−54.414 (*p*<0.001) 583 (*p*<0.001) −24.413	3.84**
Si	20.02	3.416	13.06	3.374	16.54	4.863	−29.074 (*p*<0.001) 12,205 (*p*<0.001) −20.875	2.05**
Nt	35.24	3.495	24.25	3.463	29.75	6.503	−44.77 (*p*<0.001) 2008 (*p*<0.001) −23.957	3.158**
Cs	14.82	3.086	11.95	3.047	13.39	3.385	−13.296 (*p*<0.001) 37,248.5 (*p*<0.001) −13.294	0.935*
Su	23.26	3.174	14.26	3.167	5.504	5.504	−40.227 (*p*<0.001) 3466 (*p*<0.001) −23.519	2.838**
Be	12.56	3.046	11.5	3.013	12.03	3.074	−4.959 (*p*<0.001) 65,422 (*p*<0.001) −4.694	0.35*

Note: K= Inconsistencies; L= Lies; F= Fraud; Si= Simulation; Nt= Neurasthenia; Cs= Substance Use; Su= Suggestibility; Be= Thrill-Seeking; M= Mean; SD= Standard Deviation; **= large effects; *= medium effects. Cohen's d was corrected using Hedge's g.

Table 3
Descriptive statistics of MMSI-2 scales. (Continuation Table 2).

Scales	Believers		Non-believers		Complete sample		*t test* *U test* *Z test*	*Cohen's d*
	M	SD	M	SD	M	SD		
Hi	43.59	2.903	32.62	2.884	38.11	6.206	−53.779 (*p*<0.001) 811.5 (*p*<0.001) −24.339	3.791**
Ez	35.67	4.104	23.74	4.025	29.7	7.222	−41.633 (*p*<0.001) 3279.5 (*p*<0.001) −23.565	2.935**
Pa	28.35	3.203	19.37	3.165	23.86	5.504	−39.962 (*p*<0.001) 4134 (*p*<0.001) −23.328	2.82**
Na	34.36	3.415	24.4	3.371	29.38	6.027	−41.62 (*p*<0.001) 3238.5 (*p*<0.001) −23.591	2.935**
Pva	32.11	3.805	20.14	3.778	26.13	7.088	−44.772 (*p*<0.001) 2040.5 (*p*<0.001) −23.943	3.157**
Pt	25.49	4.007	16.5	3.997	20.99	6.019	−31.848 (*p*<0.001) 9249.5 (*p*<0.001) −21.757	2.246**
Po	24.07	3.921	16.14	3.845	20.11	5.547	−28.925 (*p*<0.001) 12,327 (*p*<0.001) −20.825	2.042**
Pc	20.03	3.903	16.08	3.86	18.06	4.354	−14.428 (*p*<0.001) 38,352 (*p*<0.001) −12.923	1.017**

Note: *Hi= Histrionism; Ez= Schizotypy; Pa= Paranoia; Na= Narcissism; Pva= Anomalous Visual/Auditory Phenomena; Pt= Anomalous Tactile Phenomena; Po= Anomalous Olfactory Phenomena; Pc= Anomalous Synesthetic Phenomena; M= Mean; SD= Standard Deviation; **= large effects.* Cohen's d was corrected using Hedge's *g*.

Tucker-Lewis coefficient (TLI, threshold= >0.95); incremental fit index (IFI, threshold= >0.95); Relative fit index (RFI, threshold= >0.95); Normed fit index (NFI, threshold=>0.95); adjusted goodness of fit index (AGFI, threshold= >0.9). It should be noted that the *Chi square* statistic is highly sensitive to the sample size, and the probability of significance could be altered by simply manipulating the size of the groups (e.g., Brown, 2015). Therefore, by itself, *Chi Square* could not be interpreted. These results suggested that invariance analysis could be a good option.

When fit indices were applied to the total set assuming the constraints of the different invariance models (see Table 5), their values were satisfactorily high, and the CFI had a variation lower than 0.01. The *chi-square* statistic also showed non-significant variations (Δp *> 0.05). Unlike the analyses of a single group, in nested designs, this statis-

Table 4
Exploratory factor analysis.

Scales	Believers (N= 402)				Non-believers (N= 402)				Ordinal alphas*
	CPT	APP	IMA	ASC	CPT	APP	IMA	ASC	
Ez	0.845				0.847				0.997
Su	0.838				0.829				0.995
Pa	0.838				0.829				0.991
Na	0.832				0.831				0.997
Be	0.803				0.820				0.865
Hi	0.79				0.782				0.993
Pt		0.897				0.895			0.988
Pva		0.867				0.869			0.996
Pc		0.864				0.878			0.994
Po		0.795				0.813			0.996
F			0.684		0.632		0.670		0.993
Si			0.668		0.685		0.673		0.996
K			0.635		0.624		0.624		0.973
L			0.609		0.619		0.669		0.994
Cs				0.933				0.947	0.878
Nt				0.831				0.814	0.997
% var.	38.37	21.088	15.346	10.552	39.525	21.47	16.012	10.904	-

Note: *K= Inconsistencies; L= Lies; F= Fraud; Si= Simulation; Nt= Neurasthenia;*
Cs= Substance Use; Su= Suggestibility; Be= Thrill-Seeking; Hi= Histrionism;
Ez= Schizotypy; Pa= Paranoia; Na= Narcissism; Pva= Anomalous Visual/Auditory Phenomena; Pt= Anomalous Tactile Phenomena; Po= Anomalous Olfactory Phenomena; Pc= Anomalous Synesthetic Phenomena; CPT= Clinical Personality Tendencies; APP= Anomalous Perceived Phenomena; IMA= Incoherent Manipulations; ASC= Altered States of Consciousness; %var.= explained variance.
* Ordinal alphas came from Escolà-Gascón (2020) report.

Table 5
Model fit indices of the MMSI-2 internal structure and multi-group analysis.

Indices	Initial models		Multi-group models			
	B	NB	*Configural invariance*	*Weak factorial invariance*	*Strong factorial invariance*	*Strict Factorial invariance*
χ^2	211.937	236.912	448.849	450.25	450.408	575.966
p	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$\Delta\chi^2$	-	-	-	↑1.401	↑0.158	↑125.558
Δp^{**}	-	-	-	N.S.	N.S.	<0.0001
X^2/df	2.141	2.393	2.267	2.144	2.047	2.451
RMSEA	0.053 (0.043-0.063)	0.059 (0.049-0.069)	0.04 (0.035-0.045)	0.038 (0.033-0.041)	0.036 (0.031-0.041)	0.043 (0.038-0.047)
CFI	0.982	0.981	0.982	0.982	0.983	0.975
ΔCFI	-	-	-	0.000	↑0.001	↓0.008
TLI	0.979	0.977	0.978	0.98	0.982	0.974
IFI	0.983	0.981	0.982	0.982	0.983	0.975
RFI	0.961	0.961	0.961	0.963	0.965	0.958
NFI	0.968	0.968	0.968	0.968	0.968	0.958
AGFI	0.914	0.903	0.908	0.913	0.917	0.905

Note: B= Believers group; NB= Non-believers group; RMSEA= Root mean square error of approximation; CFI= Comparative fit index; TLI= Tucker-Lewis index; IFI= Incremental fit index; RFI= Relative fit index; NFI= Normed fit index; AGFI= Adjusted goodness of fit index; $\Delta\chi 2$= Increase in the Chi Square coefficient; Δp^{**} = Probability that the increase in Chi Square does not differ from the previous model; ΔCFI= Increase in the CFI index.

tic is also used to examine the probability that the progressive increase in the *Chi Square* itself did not differ from the initial *Chi Square* values (i.e., those belonging to the unconstrained model). If the increments present probabilities lower than 0.05 (and consequently were significant), then it would not be possible to assume factorial invariance. This would imply that the items and scales of the MMSI would not measure the same attribute in the two groups. Table 5 indicates that the increases in *Chi Square* did not yield significant probability values under the 'strong factorial (scalar) invariance' assumption. In contrast, probability did become significant under the 'strict factorial (residual) invariance' assumption. Although the ideal would have been to accept residual invariance, this situation is the most common in this type of psychometric model, and it is not necessary to assume the strictest invariance model to accept the construct validity of the test analyzed (e.g., Byrne, 2014; Little, 2013). Therefore, according to this idea, the theoretical model of the MMSI, the measurements and the content that evaluate its scales have the same meaning in both the group of believers and the group of non-believers, offering satisfactory construct validity and a good fit.

However, although the scales can have the same meaning, it does not mean that the scores of both groups are the same. Tables 2 and 3 show that the means between both groups were different. These results and the accepted invariance model allow us to examine whether the *latent means* (those belonging to the variables IMA, ASC, APP and CPT) are equal or differ from "0". This analysis can be done in several ways. The most accurate and appropriate for this type of design consists of *fixing the means of the reference group at zero.* The other means are estimated freely. In this case, the reference group is that of non-believers. In reality, the means freely estimated in the group of believers are not the empirical means of the second-order factors of the MMSI. They are average values that reflect the number of units the scores of the group of believers vary with respect to the reference group. This is the main difference with

Table 6
Latent means, descriptive statistics and reliability coefficients.

G	LV	*Means* estimates (*p*)	S.E.	*OM*	*SD*	*Hypothesis Contrast Tests*	*ω*	*α*
Believers (N=402)	CPT	9.09 *p*<0.0001	0.214	177.79	18.211	-	0.959	0.963
	IMA	7.932 *p*<0.0001	0.193	138.84	11.133	-	0.931	0.925
	ASC	10.867 *p*<0.0001	0.252	50.06	6.321	-	0.916	0.912
	APP	5.819 *p*<0.0001	0.209	101.7	14.617	-	0.952	0.952
Non-believers (N=402)	CPT	0	-	125.89	18.152	-	0.963	0.963
	IMA	0	-	95.8	11.118	-	0.957	0.955
	ASC	0	-	36.2	6.222	-	0.901	0.901
	APP	0	-	68.87	14.588	-	0.958	0.958
Complete sample (N=804)	CPT	-	-	151.84	31.693	*t test*= -40.47* *U test*= 3552* *Z test*= -23.465* *Cohen's d*= 2.854**	0.97	0.965
	IMA	-	-	117.32	24.233	*t test*= -54.843* *U test*= 569* *Z test*= -24.377* *Cohen's d*= 3.868**	0.963	0.892
	ASC	-	-	43.13	9.647	*t test*= -31.335* *U test*= 9361.5* *Z test*= -21.716 *Cohen's d*= 2.209**	0.866	0.770
	APP	-	-	85.29	21.974	*t test*= -31.880* *U test*= 8915* *Z test*= -21.838* *Cohen's d*= 2.248**	0.966	0.957

Note: *G= groups; CPT= Clinical Personality Tendencies; APP= Anomalous Perceived Phenomena; IMA= Incoherent Manipulations ASC= Altered States of Consciousness (ASC); OM= observed means; SD= standard deviation; **= large effects; ω= McDonald's Omega; α= Cronbach's Alpha.* Cohen's d was corrected using Hedge's *g*.

respect to comparison tests of groups based on the contrast of means or ranks (see Brown, 2015). Therefore, constraining the means associated with the latent variables of the non-believer group to "zero", the average dispersion values shown in Table 6 were obtained. The *p*-value of this table indicates the probability that the scores of the Believers group are distributed according to the distribution given by the scores of the reference group. All the mean estimates obtained significant probability values (<0.0001).

According to the results in Table 6, this means that the scores of the believing subjects are significantly higher than the scores of the group of non-believers. This result was confirmed through statistical hypothesis testing generated for the empirical means of each group. These data are also shown in Table 6.

3.3. Simultaneous regression models between latent variables

Figs. 2 and 3 show that the standardized covariances between the latent variables of the model were low. It was asked if this situation could change using the total sample (N = 804) and using the simultaneous regression models. This is justified by three main reasons: (1) in other studies, the macro-factors of the MMSI were extracted from oblique factorial solutions, allowing the possibility that the macro-factors were correlated (see Escolà-Gascón, 2020; Escolà-Gascón and Gallifa, 2020). (2) Taking into account the results of the factorial invariance, contrasting this possibility would also facilitate knowing if these belief systems modulate the covariances between the macro-factors. (3) In addition, according to Abad et al. (2015), in the behavioral sciences, orthogonal solutions are unusual, and it is most likely that there would be intercorrelations between the constructs. Therefore, to caution against possible intercorrelations between macro-factors and following the recommendations of Arribas (2011), it was decided to contrast the structural model through simultaneous regressions, as shown in Fig. 4. The 4 macro-factors defined in the previous Figs. were chosen, and their predictive value was examined, especially on the APP variable. To determine the fixing of the

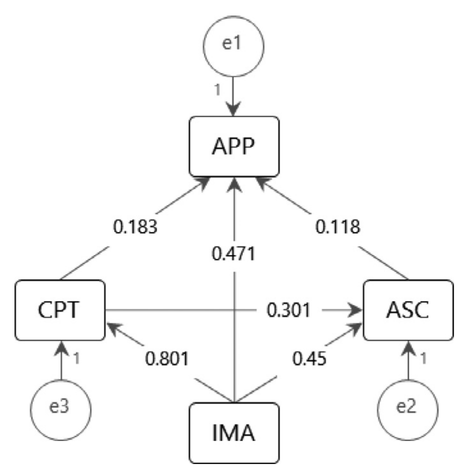

Fig. 4. Simultaneous regression model. Regression weights are provided.

effects, previous research that verified significant correlations between the latent variables was also taken into account.

As was suspected, when the two groups are merged, the linear correlations between the 4 factors increase significantly. Given that the regressions had been developed using the *maximum likelihood* method, the following fit indices were calculated: $\chi^2 = 8.424$ ($p = 0.004$); *normed* $\chi^2 = 8.424$; CFI = 0.996; TLI = 0.979; IFI = 0.996; RFI = 0.976; NFI = 0.996; AGFI = 0.948; and RMSEA = 0.09 (0.04-0.16).

The IMA, CPT and ASC variables explained a total of 51.2% of the variance of APP. IMA explained 64.2% of the variance of CPT. Both IMA and CPT explained 50.9% of the ASC variance. This contrast shows that IMA, CPT and ASC predict the anomalous phenomena evaluated by APP. Table 7 shows the matrix of linear correlations between macro-factors.

Table 7
Matrix correlations between latent variables.

	IMA	CPT	ASC	APP
IMA	-			
CPT	0.823*	-		
ASC	0.7*	0.677*	-	
APP	0.706*	0.654*	0.574*	-

Note: CPT= Clinical Personality Tendencies; APP= Anomalous Perceived Phenomena; IMA= Incoherent Manipulations ASC= Altered States of Consciousness (ASC); *p<0.001.

These results suggest that the variable beliefs in existence in the paranormal could have moderating effects on the relationship between macro-factors. When this variable is isolated and recorded as a constant (establishing the two groups and imposing invariance constraints), the macro-factors lose predictive power, and the solution tends to be orthogonal.

3.4. Reliability analysis

Cronbach's Alpha and *McDonald's Omega* reliability coefficients also shown in Table 6. The values of these coefficients will be *acceptable* from 0.7 and *excellent* when they are greater than 0.9 (e.g., George and Mallery, 2003; Trizano-Hermosilla and Alvarado, 2016). Both the *omega* and *alpha* indices were calculated for each of the groups (believers and non-believers), as well as for the total set of the sample. All of them were acceptable and excellent in several cases, so it was not necessary to eliminate any scale to optimize these indices. Likewise, the lowest coefficients (especially that of the ASC dimension) can be associated with the small number of scales that are grouped. In general, the results of these indices allow us to satisfactorily accept the reliability of internal consistency in MMSI-2 scores.

4. Discussion

First, the validity and reliability of the MMSI-2 was tested using the statistical techniques of the structural equations, the analysis of the 'factorial invariance' and the reliability coefficients applied. The three designs showed results supporting the construct validity of the MMSI-2 and the reliability associated with its scores. Second, the initial objective of the study—based on the validation of a structural and measurement model with respect to anomalous phenomena—also had the purpose of contrasting the empirical relationship between the variables that, according to the published background, correlated with the anomalous phenomena. In this case, through the simultaneous regression models, it was possible to conclude that the 4 latent variables (CPT, ASC, APP and IMA) were positively correlated with each other when both groups were merged. Finally, the third objective was to examine the variability or change in scores between the group of believers and the group of non-believers. This was carried out through comparison of means tests and the 'latent means' analysis of the factorial model. All results supported the hypothesis that believers in the existence of the paranormal scored significantly and in most scales above those of non-believers. Each of these findings in the results raises questions and new hypotheses related to the problems formulated in the theoretical framework. Two questions of interest for clinical psychology and applied psychology in the forensic realm can be posed: (1) why do believing subjects score systematically higher than non-believing subjects? Moreover, are systems of meaning truly responsible for these differences? (2) What information or utility do the MMSI-2 scales—especially the CPT, ASC and IMA macro-factors—have in the evaluation of those behaviors that are "apparently" without scientific explanation or that are extremely uncertain and divergent with the clinical discourse?

One of the most important points was whether subjects who believe in the paranormal interpret and conceive anomalous phenomena in the same way as non-believers. The phenomenological model originally proposed by Jaspers (1993) and Irwin (2009) postulates that one can observe a change in 'systems of meanings' and 'causal attributions' between believing and non-believing subjects. According to this approach, believing subjects would be more vulnerable to experiencing anomalous phenomena because they possess cognitive systems capable of interpreting any situation under an attribution of paranormal cause (e.g., Irwin et al., 2013). This idea can be related to the results obtained in the comparison of means test (see Table 2). In all the contrasts, the means of the believing subjects were higher than the means of the non-believing subjects, especially in the scales grouped in APP. This convergence and compatibility with the postulates of Irwin et al. (2013) is also reflected in the analyses of 'latent means'. However, there is the question of whether these differences in the systems of meanings are truly responsible for one group to score higher than the other. These differences cannot be negated qualitatively since the contents and categories of each system of meanings are different in each group. However, it can be questioned where they come from. It would seem logical to expect the rejection of the 'strong factorial invariance' assumption if one starts from the basis that believers interpret anomalous phenomena differently than non-believers. However, the applied structural equation models present fit indices with satisfactory results and non-significant variations when imposing invariance constraints. The statistical interpretation derived from these results supports the possibility that, in effect, both groups understand and the contents of each scale in the same way. This calls into question the reasons why believing subjects systematically score higher than others. The qualitative and categorical differences between the cognitive systems of believers and non-believers are present and obvious (see French and Stone, 2014). However, these results indicate that these differences do not cause the scores to be higher in the group of believers. Even imposing these invariance constraints, the 'latent means' of the believers deviate between 5 and almost 11 units with respect to those of the group of non-believers.

This does not contradict the hypothesis of Irwin (2009), but it does warn that other psychological mechanisms could intervene and generate these differences. Taking into account the studies published by several professionals, it is possible that the variables responsible for these differences are related not only to beliefs but also to *cognitive* or *causal learning* mechanisms (e.g., Barberia et al., 2013; Barberia et al., 2018; Groome et al., 2019).

Regarding the psychological value provided by macro-factors in the psychometric evaluation of anomalous phenomena, two types of interpretations can be identified: on the one hand, the results observed in Figs. 2 and 3 and in the analysis of factorial invariance can be used. On the other hand, the simultaneous equations models that relate and allow predicting the APP variable with a total weight of 51.2% can also be used. While the first possibility offers an 'orthogonal' interpretation, the second is based on the relationship and prediction among the 4 factors, so it offers a more 'oblique' view. The CFAs applied to the two groups suggest that APP, ASC, IMA and CPT are not significantly related to each other. They present low correlations (the most relevant ones fluctuate between 0.1 and 0.3) and do not allow estimating the coefficient of determination (R^2). This does not preclude the interpretation of the anomalous phenomena evaluated by the MMSI, but it does make it difficult because it would not be possible to distinguish the psychological antecedents from the anomalous phenomena. However, this only occurs when subjects are differentiated according to whether they believe in the paranormal. When contrasting the null hypothesis of independence and the predictive value of macro-factors in the total sample of the study, the results are completely different. Fig. 4 clearly shows that the factors are interrelated. This antagonistic change between some models and others can be explained by two main reasons, but both are related to the criteria used for the distribution of the two groups.

First, as mentioned in the section on results 3.3., it is possible that the grouping variable (i.e., the belief in the existence of the paranormal) intervenes as *a moderating variable* (and not necessarily a *mediator*) in

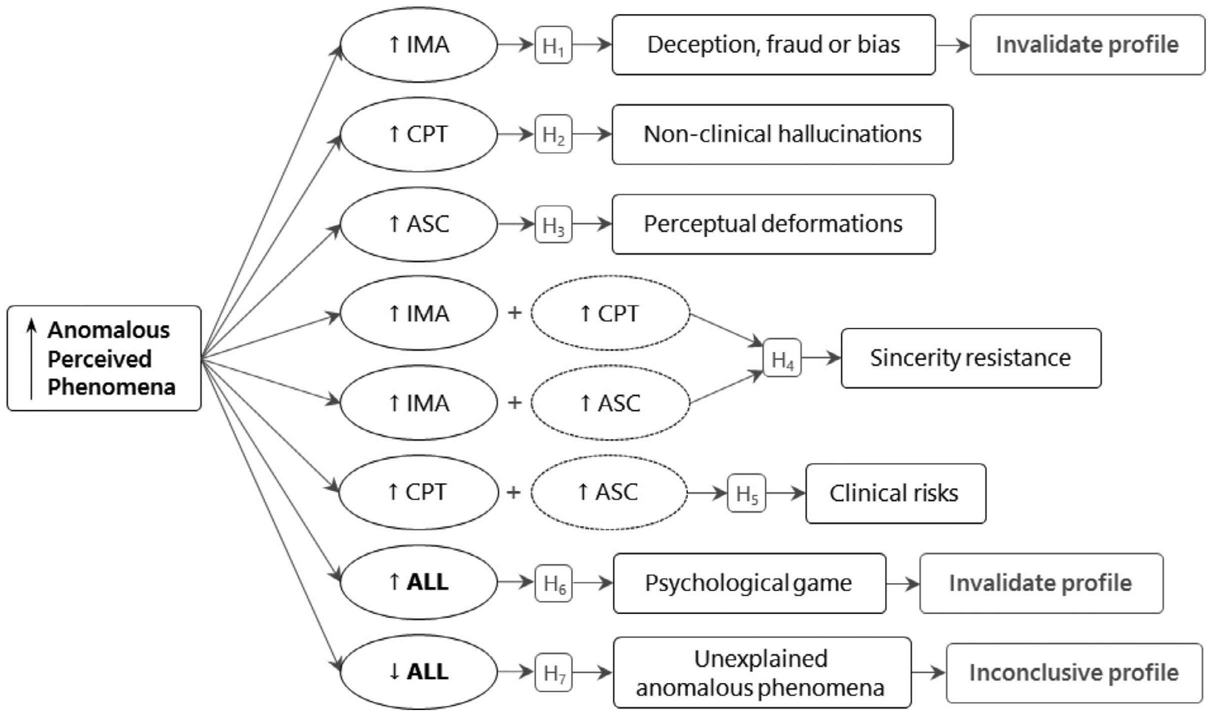

Fig. 5. Diagram of hypothetical formulations based on MMSI-2 scores.

the relationship of IMA, ASC and CPT with APP. This allows noting the possible interaction between this moderator variable and the IMA, ASC and CPT macro-factors. This is supported by the results in the 'latent means' analyses and the comparison of means.

Second, in addition to this first possibility, one must also take into account how both groups have been formed and with what characteristics. The most obvious is that the quantities for each of the values relating to the sociodemographic variables were matched. This may have favored homogeneity in the covariance between the scores of one group and the scores in the other, which optimizes the 'factorial invariance' analysis. However, it is possible that this has also penalized the covariability between the latent variables. Unless the factorial solution used is based on 'oblique' extraction criteria, which is not the case because the EFAs in this report have not been manipulated obliquely, this homogeneity in the covariances decreases as factors of a higher order are extracted (see Mulaik, 2018). In any of the cases, Abad et al. (2015) warn that completely 'orthogonal' solutions are not common in this type of analysis and, therefore, recommend replicating the intercorrelations and the predictive value among the latent variables of the model.

If the relationship of the four latent variables is based on Fig. 4, then one could *hypothesize* mechanisms or criteria to differentiate the explanation and classification of anomalous phenomena. The diagram in Fig. 5 was developed according to the contents grouped in each macro-factor. It summarizes the possible hypothetical interpretations that should be investigated and contrasted in future research.

Fig. 5 presents only hypothetical associations. The hypotheses in this diagram do not come directly from the results of this research. However, based on an evaluation without discriminating belief systems and taking into account the contents belonging to each scale, it is possible to deduce the hypotheses presented in Fig. 5.

Considering the dilemma generated by the investigation of the 'psi' phenomena and their impact on psychological evaluation, the MMSI-2 scales offer the possibility of formulating scientific hypotheses about the etiology and classification of anomalous phenomena. While conventional stereotypes tend to incur (though not always) in the 'Aristotelian fallacy of verification of the consequent', including statistical decisions (see Pardo and Román, 2013), MMSI-2 represents a resource to prevent this type of error. It attempts to substantiate the observations and suspicions of the professional-researcher working in the field of mental health. It is as erroneous to assert that 'psi' phenomena exist to deny their possible existence. It is erroneous to accept that anomalous phenomena have a "parapsychological" origin, as all of them are hallucinations related to psychosis (e.g., French and Stone, 2014). Academic research related to cognitive and perceptual processes should be based on the application of scientific methodology through the testing of empirical indicators (*scientific empiricism*) and should not be limited to the exclusive use of *scientific rationalism* (e.g., Carter, 2012). Precisely, one way to combat these argumentative and fallacious errors rests on the psychometric development of scientific evaluation protocols, such as the MMSI-2. It should be noted that the MMSI does not verify or confirm the causal antecedents of anomalous phenomena (APP). However, Figs. 4 and 5 at least empirically ground the explanatory psychological hypotheses and would help in the prevention of *type I* and *II errors*. Given that the criteria in Fig. 5 are designed to be applied both in the statistical study and at the individual level, new research is needed to test the predictive validity of Fig. 5 and the quality of the scales, especially with regard to cut-off scores. This is discussed again in the following paragraphs.

One of the most obvious limitations of this research is that the CFAs have not been directly applied to the 174 items of the MMSI-2. This entailed assuming the primary scales as the observable variables of the model and the macro-factors as the respective latent variables. This decision followed the statistical model applied by Arribas (2011) in the *TEA Personality Test* (TPT). The same idea has been used in multiple self-report questionnaires (see Butcher et al., 2019; Gorsuch, 1983; Morey, 2011). The advantage it offers with respect to the conventional CFA models based on the items is that in this class of models, the analysis of the structural model is optimized. However, a demonstrated disadvantage and limitation is that accuracy (but not information) is lost with respect to the study of the metric quality of the items. In any case, as indicated by Mulaik (2018), if the items presented problems of covariance or measurement, these errors would also affect the higher order factor analysis. Given that the solutions and models presented seem acceptable in terms of goodness of fit, the metric of the items would not impair the results obtained.

As a second limitation, it can be noted that the interpretative hypotheses in Fig. 5 lack objectivity if cutoff points for each scale and latent factor of the MMSI-2 are not specified. This criticism is related to the scoring of the direct scores and the sensitivity and specificity of the test. Although Spanish scales are available to obtain guidance thresholds (see Escolà-Gascón and Gallifa, 2020) and thus make evaluative decisions, it would be advantageous to test the sensitivity and specificity of the MMSI with new samples and external evaluation criteria. This would also be related to the possibility of validating the MMSI-2 with a clinical population. The latter would allow examining cut-off points not only to discriminate what is elevated and what is not; it could also contrast the discriminative value of what is psychopathological and what is sub-clinical.

Finally, it should be noted that new psychometric research is needed to explore and replicate the relationship between primary scales, macro-factors and other variables of interest. Examples are those attributes related to the detection of lying, simulation and fraud. Although the MMSI-2 has a scale that is intended for the assessment of deliberate fraud (F scale), it would be necessary to conduct an experimental study between groups (one of simulant subjects and another of non-simulant subjects), which contrasts with the degree to which the F scale and the IMA factor are able to identify simulation behaviors. This same logic can be extrapolated with other clinical and educational variables.

It can be concluded that the MMSI-2, with its 20 total scales, represents a valid and reliable psychometric instrument for the examination of anomalous phenomena and other concomitant psychological variables. The test also shows that the psychological (and non-psychopathological) etiology of anomalous phenomena is conditioned and can be estimated hypothetically from the macro-factors 'Clinical Personality Tendencies' (CPT), 'Incoherent Manipulations' (IMA, related to fraud and lie detection) and 'Altered States of Consciousness' (ASC), all of which were examined by the MMSI-2. Therefore, the MMSI-2 can be a useful tool for evaluation and lie-detection in the subjects who report anomalous experiences. Likewise, the MMSI-2 can also be applied in the clinical scope with the purpose of discriminating if the anomalous experiences are unexplained experiences, perceptive deformations or hallucinations. In conclusion, the most relevant contribution of the MMSI-2 is being the only psychometric instrument designed in the field of psi research, that offers objective measurements to know if anomalous experiences have a psychological explanation or not.

Author contributions

The author conceived and planned the study, drafted and designed the questionnaire, collected the sample, performed the statistical analyses and wrote the present report.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics statement

Participants gave their written consent to use their anonymous data for statistical purposes. All of them were over 18 years old and voluntarily collaborated without receiving any financial compensation. The procedures were carried out in compliance with the institutional regulations of the *Ramon Llull University* and the Spanish Government *Data Protection Law 15/1999*. Similarly, all procedures adhere to the *Helsinki Declaration of 1975*, revised in 2013.

Datasets are available on request

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

The author wishes to express his gratitude to the assistants and organizations that collaborated with sample collection and data entry: to the congregation *Sisters of Immaculate Conception* (Madrid, Spain); and, the professionals *Judith García* (Barcelona, Spain), *Ramon Palau* (Barcelona, Spain), *Rosa María Gamboa* (Barcelona, Spain) and *Sister Bárbara Ávila* (Madrid, Spain).

References

Abad, F.J., Olea, J., Ponsoda, V., García, C., 2015. Medición en Ciencias Sociales y de la Salud. Editorial Síntesis.
Arribas, D., 2011. Psychometric properties of the TEA personality test: evidence of reliability and construct validity. Eur. J. Psychol. Assessment 27 (2), 121–126. doi:10.1027/1015-5759/a000057.
Barberia, I., Blanco, F., Cubillas, C.P., Matute, H., 2013. Implementation and assessment of an intervention to debias adolescents against causal illusions. PLoS One 8. doi:10.1371/journal.pone.0071303, Article e71303.
Barberia, I., Tubau, E., Matute, H., Rodríguez-Ferreiro, J., 2018. A short educational intervention diminishes causal illusions and specific paranormal beliefs in undergraduates. PLoS One 13 (1). doi:10.1371/journal.pone.0191907, Article e0191907.
Beischel, J., Boccuzzi, M., Biuso, M., Rock, A., 2015. Anomalous information reception by research mediums under blinded conditions II: Replication and extension. Explore 11, 136–142. doi:10.1016/j.explore.2015.01.001.
Beischel, J., Zingrone, N., 2015. Mental Mediumship. In: Cardeña, E., Palmer, J., Mar-cusson-Clavertz, D. (Eds.), Parapsychology: A Handbook for the 21st Century. McFarland and Company, Inc, pp. 301–313.
Bell, V., Halligan, P.W., Ellis, H.D., 2006. The cardiff anomalous perceptions scale (CAPS): a new validated measure of anomalous perceptual experience. Schizophr. Bull. 32 (2), 366–377. doi:10.1093/schbul/sbj014.
Belloch, A., Baños, R.M., Perpiñá, C., 1995. Psicopatología de la percepción y la imaginación. In: Belloch, A., Sandín, B., Ramos, F. (Eds.), Manual de psicopatología: volumen I. McGraw-Hill, pp. 188–230.
Bem, D.J., 2011. Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. J. Pers. Soc. Psychol. 100 (3), 407–425. doi:10.1037/a0021524.
Bem, D.J., Tressoldi, P., Rabeyron, T., Duggan, M., 2016. Feeling the future: a meta-analysis of 90 experiments on the anomalous anticipation of random future events. F1000research 4. doi:10.12688/f1000research.7177.2, Article 1188.
Bobrow, R.S., 1983. A choice to die. Psychology Today 17 (6), 70–72.
Bobrow, R.S., 2003. Paranormal phenomena in the medical literature sufficient smoke to warrant a search for fire. Med. Hypotheses 60 (6), 864–868. doi:10.1016/S0306-9877(03)00066-5.
Brown, T.A., 2015. Confirmatory Factor Analysis for Applied Research. The Guilford Press.
Bunge, M. (2013). *La Ciencia, su Método y su Filosofía*. Laetoli, S.L.
Butcher, J.N., Graham, J.R., Ben-Porath, Y.S., Tellegen, A.W., Dahlstrom, W.G., Kaemmer, B., 2019. MMPI-2. Inventario Multifásico de Personalidad de Minnesota -2 (manual). TEA Ediciones, S.A.U.
Butcher, J.N., Graham, J.R., Ben-Porath, Y.S., 1995. Methodological problems and issues in MMPI, MMPI-2, and MMPI-A research. Psychol. Assess. 7 (3), 320–329. doi:10.1037/1040-3590.7.3.320.
Byrne, B.M., 2014. Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming. Psychology Press.
Byrne, B.M., Shavelson, R.J., Muthén, B., 1989. Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. Psychol. Bull. 105 (3), 456–466. doi:10.1037/0033-2909.105.3.456.
Cameron, R., 2016. The paranormal as an unhelpful concept in psychotherapy and counselling research. Eur. J. Psychotherapy Counsell. 18 (2), 142–155. doi:10.1080/13642537.2016.1170060.
Carter, C., 2012. Science and Psychic Phenomena: The Fall of the House of Skeptics. Inner Traditions.
Chapman, L.J., Chapman, J.P., Raulin, M.L., 1978. Body-image aberration in schizophrenia. J. Abnorm. Psychol. 87 (4), 399–407. doi:10.1037/0021-843X.87.4.399.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Academic Press.
David, A.S., 2010. Why we need more debate on whether psychotic symptoms lie on a continuum with normality. Psychol. Med. 40 (12), 1935–1942. doi:10.1017/s0033291710000188.

Drinkwater, K., Dagnall, N., Grogan, S., Riley, V., 2017. Understanding the unknown: a thematic analysis of subjective paranormal experiences. Austr. J. Parapsychol. 17 (1), 23–46.
Escolà-Gascón, Á, 2020. Researching unexplained phenomena: empirical-statistical validity and reliability of the Multivariable Multiaxial Suggestibility Inventory-2 (MMSI-2). Heliyon 6 (7). doi:10.1016/j.heliyon.2020.e04291, Article e04291.
Escolà-Gascón, Á., Gallifa, J., 2020. Psychology of anomalous experiences: psychometric properties of the Multivariable Multiaxial Suggestibility Inventory -2 Reduced (MMSI-2-R). Anuario de Psicología/The UB J. Psychol. Advance online publication. Article 28189.
Eysenck, H.J., Sargent, C., 1982. Los misterios de lo Paranormal. Editorial Planeta, S.A..
Fekih-Romdhane, F., Sassi, H., Ennaifer, S., Tira, S., Cheour, M., 2020. Prevalence and correlates of psychotic like experiences in a large community sample of young adults in Tunisia. Community Ment. Health J. doi:10.1007/s10597-019-00542-1, Advance online publication.
Fernández-Ballesteros, R., 2011. Evaluación Psicológica: Conceptos, Métodos y Estudio de Casos. Ediciones Pirámide.
Fleiss, J.L., Levin, B., Cho-Paik, M., 2003. Statistical Methods for Rates and Proportions. Wiley Inter-Science, INC.
Font, J., 2016. Religión, Psicopatología y Salud Mental: Introducción a la Psicología de las Experiencias Religiosas y de las Creencias. Ediciones Paidós y Fundació Vidal i Barraquer.
French, C.C., Stone, A., 2014. Anomalistic Psychology: Exploring Paranormal Belief and Experience. Red Globe Press, Inc.
George, D., Mallery, P., 2003. SPSS for Windows Step by step: A Simple Guide and Reference. 11.0. Update, 4th Ed. Allyn and Bacon.
Gorsuch, R.L., 1983. Factor Analysis. Lawrence Erlbaum.
Groome, D., Eysenck, M., Law, R., 2019. The Psychology of the Paranormal. Taylor and Francis Group, Inc, Routledge.
Groth-Marnat, G., 2009. Handbook of Psychological Assessment. John Wiley and Sons, Ltd.
Harary, K., 2006. Experiencias paranormales: sugerencias para los psicólogos clínicos. In: Parra, A. (Ed.), Psicología de las Experiencias Paranormales. Librería Akadia Editorial, pp. 57–80.
Heotis, E., 2019. Mental imagery and its limitations. Curr. Opin. Neurol. Sci. 3 (2), 648–651.
Honorton, C., 1985. Meta-analysis of psi ganzfeld research: a response to Hyman. J. Parapsychol. 49, 51–91.
Houran, J., Laythe, B., O'Keeffe, C., Dagnall, N., Drinkwater, K., Lange, R., 2019. Quantifying the phenomenology of ghostly episodes: part I - need for a standard operationalization. J. Parapsychol. 83 (1), 25–46. http://doi.org/10.30891/jopar.2019.01.03.
Irwin, H.J., 1993. Belief in the paranormal: a review of the empirical literature. J. Am. Soc. Psychical Res. 87, 1–39.
Irwin, H.J., 2000. Belief in the paranormal and a sense of control over life. Eur. J. Parapsychol. 15, 68–78.
Irwin, H.J., 2003. Paranormal beliefs and the maintenance of assumptive world views. J. Soc. Psychical Res. 67, 18–25.
Irwin, H.J., 2009. The Psychology of Paranormal Belief. University of Hertfordshire Press.
Irwin, H.J., Dagnall, N., Drinkwater, K., 2013. Parapsychological experience as anomalous experience plus paranormal attribution: a questionnaire based on a new approach to measurement. J. Parapsychol. 77, 39–53.
Irwin, H.J., Watt, C., 2007. An Introduction to Parapsychology. McFarland and Company, Inc.
Jaccard, J., Wan, C.K., 1996. LISREL Approaches to Interaction Effects in Multiple Regression. Sage Publications, Ltd.
Jaspers, C., 1993. Psicopatología General. Editorial Beta.
Jinks, T., 2019. Psychological Perspectives on Reality, Consciousness and Paranormal Experience. Springer Nature, Inc.
Johns, L.C., van Os, J., 2001. The continuity of psychotic experiences in the general population. Clin. Psychol. Rev. 21 (8), 1125–1141. doi:10.1016/S0272-7358(01)00103-9.
Kelly, E.W., Arcangel, D., 2011. An investigation of mediums who claim to give information about deceased persons. J. Nerv. Ment. Dis. 199, 11–17. doi:10.1097/NMD.0b013e31820439da.
Kelly, R., Shoulder, C., Bell, V., 2020. Assessment in psychosis. In: Badcock, J., Paulik, G. (Eds.), A Clinical Introduction to Psychosis: Foundations for Clinical Psychologists and Neuropsychologists. Academic Press, ELSEVIER, pp. 135–152. doi:10.1016/B978-0-12-815012-2.00006-7.
Kuhn, G., Olson, J., Raz, A., 2016. Editorial: the psychology of magic and the magic of psychology. Front. Psychol. 7. doi:10.3389/fpsyg.2016.01358, Article 1358.
Kline, P., 2013. Handbook of Psychological Testing. Taylor and Francis Group, Routledge.
Lange, R., Ross, R., Dagnall, N., Irwin, H., Houran, J., Drinkwater, K., 2019. Anomalous experiences and paranormal attributions: psychometric challenges in studying their measurement and relationship. Psychol. Conscious.. Advance online publication http://dx.doi.org/10.1037/cns0000187 .
Lawrence, T., 2016. The client, the therapist and the paranormal: a response to the special edition on psychotherapy and the paranormal. Eur. J. Psychotherapy Counsel. 18 (2), 179–192. doi:10.1080/13642537.2016.1172099.
Leonard, C.A., Williams, R.J., 2019. Fallacious beliefs: gambling specific and belief in the paranormal. Can. J. Behav. Sci. / Revue canadienne des sciences du comportement 51 (1), 1–11. doi:10.1037/cbs0000113.
Little, T.D., 2013. Longitudinal Structural Equation Modeling. Guilford Press.
Nordgaard, J., Buch-Pedersen, M., Hastrup, L., Haahr, U., Simonsen, E., 2019. Measuring psychotic-like experiences in the general population. Psychopathology 52 (4), 240–247. doi:10.1159/000502048.
Maher, M.C., 1999. Riding the waves in search of the particles: a modern study of ghosts and apparitions. J. Parapsychol. 63, 47–80.
Maher, M.C., 2000. Quantitative investigation of the General Wayne Inn. J. Parapsychol. 64, 365–390.
Maher, M.C., 2015. Ghosts and poltergeists: an eternal enigma. In: Cardeña, E., Palmer, J., Marcusson-Clavertz, D. (Eds.), Parapsychology: A Handbook for the 21st Century. McFarland and Company, Inc, pp. 327–340.
Maher, M.C., Hansen, G.P., 1992. Quantitative investigation of a reported haunting using several detection techniques. J. Am. Soc. Psychical Res. 86, 347–374.
Maher, M.C., Hansen, G.P., 1995. Quantitative investigation of a haunted castle in New Jersey. J. Am. Soc. Psychical Res. 89, 19–50.
Mason, O., Claridge, G., 2006. The Oxford-liverpool inventory of feelings and experiences (O-LIFE): further description and extended norms. Schizophr. Res. 82 (2), 203–211. doi:10.1016/j.schres.2005.12.845.
Mayer, G., 2017. What about parapsychology and anomalistics? Results of a WGFP and GFA member survey. J. Soc. Psychical Res. 81 (4), 209–227.
McDonald, R.P., 1999. Factor Analysis and Related Methods. Erlbaum.
Miller, L.A., Lovler, R.L., 2020. Foundations of Psychological Testing. SAGE Publications, Inc.
Morey, L.C., 2011. PAI. Inventario de Evaluación de la Personalidad (manual). TEA Ediciones, S.A.U..
Mossbridge, J., Tressoldi, P., Utts, J., 2012. Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. Front. Psychol. 3. doi:10.3389/fpsyg.2012.00390, Article 390.
Mulaik, S.A., 2018. Fundamentals of Common Factor Analysis. In: Irwing, P., Booth, T., Hughes, D.J. (Eds.), The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development. John Wiley and Sons, Ltd, pp. 211–252.
O'Keeffe, C., Wiseman, R., 2005. Testing alleged mediumship: Methods and results. Br. J. Psychol. 96, 165–179. doi:10.1348/000712605X36361.
Oliver, D., Reilly, T., Baccaredda Boy, O., Petros, N., Davies, C., Borgwardt, S., McGuire, P., Fusar-Poli, P., 2019. What causes the onset of psychosis in individuals at clinical high risk? A meta-analysis of risk and protective factors. Schizophr. Bull. 46 (1), 110–120. doi:10.1093/schbul/sbz039.
Pardo, A., Román, M., 2013. Reflections on the Baron and Kenny model of statistical mediation. Anales de Psicología / Ann. Psychol. 29 (2), 614–623. doi:10.6018/analesps.29.2.139241.
Parker, A., 2006. Experiencias paranormales: ¿Normalidad o anormalidad? In: Parra, A. (Ed.) Psicología de las experiencias paranormales. Librería Akadia Editorial, pp. 17–33.
Parkinson, G., 2019. Everyday magic: surrealism and parapsychology. In: Poller, J. (Ed.), Altered Consciousness in the Twentieth Century. Taylor and Francis Group, Routledge, pp. 78–96.
Pasricha, S., 2011. Relevance of parapsychology in psychiatric practice. Ind. J. Psych. 53 (1), 4–8. doi:10.4103/0019-5545.75544.
Reber, A.S., Alcock, J.E., 2019. Searching for the impossible: parapsychology's elusive quest. Am. Psychol. doi:10.1037/amp0000486, Advance online publication.
Robertson, T.J., Roy, A.E., 2001. A preliminary study of the acceptance by non- recipients of medium's statement to recipients. J. Soc. Psych. Res. 65, 91–106.
Robertson, T.J., Roy, A.E., 2004. Results of the application of the Robertson-Roy protocol to a series of experiments with mediums and participants. J. Soc. Psych. Res. 68 (1), 18–34.
Roy, A.E., Robertson, T.J., 2001. A double-blind procedure for assessing the relevance of a medium's statements to a recipient. J. Soc. Psych. Res. 65, 161–174.
Schwartz, G.E.R., Russek, L.G.S., 2001. Evidence of anomalous information retrieval between two mediums: telepathy, network memory resonance, and continuance of consciousness. J. Soc. Psych. Res. 65, 257–275.
Shapiro, D.I., Li, H., Kline, E.R., Niznikiewicz, M.A., 2019. Assessment of risk for psychosis. In: Li, H., Shapiro, D., Seidman, L. (Eds.), Handbook of Attenuated Psychosis Syndrome Across Cultures. Springer, Inc, pp. 7–40.
Shermer, M., 2011. The Believing Brain: From ghosts and Gods to Politics and Conspiracies – How we Construct Beliefs and Reinforce Them as Truths. Times Books.
Sirvent, C., Herrero, J., Moral, M.V., Rodríguez, F.J., 2019. Evaluation of self-deception: Factorial structure, reliability and validity of the SDQ-12 (self-deception questionnaire). PLoS One 14 (1), e0210815. doi:10.1371/journal.pone.0210815.
Stefanis, N.C., Hanssen, M., Smirnis, N.K., Avramopoulos, D.A., Evdokimidis, I.K., Stefanis, C.N., van Os, J., 2002. Evidence that three dimensions of psychosis have a distribution in the general population. Psychol. Med. 32 (2), 347–358. doi:10.1017/S0033291701005141.
Project, The Jamovi, 2019. Jamovi (Version 1.0) [Computer Software]. Retrieved from https://www.jamovi.org.
Tressoldi, P.E., Utts, J., 2015. Statistical guidelines for empirical studies. In: Cardeña, E., Palmer, J., Marcusson-Clavertz, D. (Eds.), Parapsychology: A Handbook for the 21st Century. McFarland and Company, Inc, pp. 83–93.
Trizano-Hermosilla, I., Alvarado, J.M., 2016. Best alternatives to cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. Front. Psychol. 7. doi:10.3389/fpsyg.2016.00769, Article 769.
Upthegrove, R., Broome, M., Caldwell, K., Ives, J., Oyebode, F., Wood, S., 2015. Understanding auditory verbal hallucinations: a systematic review of current evidence. Acta Psychiatr. Scand. 133 (5), 352–367. doi:10.1111/acps.12531.
Utts, J., 2018. An Assessment of the evidence for psychic functioning. J. Parapsychol. 82 (3), 118–146. http://doi.org/10.30891/jopar.2018S.01.10.
Vandenberg, R.J., Lance, C.E., 2000. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Org. Res. Methods 3, 4–69. doi:10.1177/109442810031002.
van Os, J., Linscott, R.J., Myin-Germeys, I., Delespaul, P., Krabbendam, L.A., 2009. A systematic review and meta-analysis of the psychosis continuum: evidence for a psy-

chosis proneness-persistence-impairment model of psychotic disorder. Psychol. Med. 39 (2), 179–195. doi:10.1017/S0033291708003814.

Wagenmakers, E., Wetzels, R., Borsboom, D., van der Maas, H., 2011. Why psychologists must change the way they analyze their data: the case of PSI: comment on bem (2011). J. Pers. Soc. Psychol. 100 (3), 426–432. doi:10.1037/a0022790.

Wahbeh, H., Yount, G., Vieten, C., Radin, D., Delorme, A., 2019. The noetic experience and belief scale: a validation and reliability study. F1000research 8 doi:10.12688/f1000research.20409.1, Article 1741.

Waugh, M.H., Hopwood, C.J., Krueger, R.F., Morey, L.C., Pincus, A.L., Wright, A.G.C., 2017. Psychological assessment with the DSM–5 alternative model for personality disorders: tradition and innovation. Profess. Psychol. 48 (2), 79–89. doi:10.1037/pro0000071.

Weiner, I.B., Greene, R.L., 2017. Handbook of Personality Assessment. John Wiley and Sons, Inc.

Wright, A., Nelson, B., Fowler, D., Greenwood, K., 2020. Perceptual biases and metacognition and their association with anomalous self experiences in first episode psychosis. Conscious. Cogn. 77. doi:10.1016/j.concog.2019.102847, Article 102847.

Wright, A., Hallquist, M.N., 2020. The Cambridge Handbook of Research Methods in Clinical Psychology. Cambridge University Press.