Journal of Parapsychology
2019, Vol. 83, No. 2, 168-192

http://doi.org/10.30891/jopar.2019.02.05

Quantifying the Phenomenology of Ghostly Episodes: Part II – A Rasch Model of Spontaneous Accounts[1]

James Houran[a], Rense Lange[a], Brian Laythe[b], Neil Dagnall[c], Kenneth Drinkwater[c], and Ciarán O'Keeffe[d]

[a] Polytechnic Institute of Management and Technology, [b] Indiana University Southeast, [c] Manchester Metropolitan University, [d] Buckinghamshire New University

Abstract. Using a sample of self-reported "spontaneous" accounts (ostensibly sincere and unprimed, *N* = 426), we calibrated a 32-item, Rasch-based "Survey of Strange Events (SSE)" to quantify the phenomenology of ghostly episodes while assessing response biases related to experients' age and gender. This inventory included psychological experiences typical of haunts, and physical manifestations common to poltergeist-like disturbances. Results supported earlier suggestions that "spontaneous" accounts have a predictable (cumulative) behavioral pattern and show a unidimensional factor structure. Further, compared to spontaneous accounts, we identified strong response biases on the SSE across four control conditions (i.e., Lifestyle, Primed, Fantasy, and Illicit). Statistical modeling successfully predicted group memberships with good accuracy, corroborating that spontaneous experiences differ systematically in certain ways from "impostors." The SSE is a robust measure of overall intensity of ghostly episodes (Rasch reliability = 0.87) and serves as a standard operationalization of specific anomalies in surveys, fieldwork studies, and investigations that code free-response data or spontaneous case material for quantitative analysis.
Keywords: ghost, haunt, phenomenology, poltergeist, psychometrics, Rasch scaling

"Repeated and intermittent ...displays are typical of haunt and poltergeist episodes. These events... involve measured or inferred physical changes such as object movements, electrical failures, or strange sounds. Reports of psychological experiences include 'odd feelings,' intelligible phrases, and sometimes the perception of human forms. Although the instances may be numerous within the lifespan of the phenomenon, the duration of a single event rarely exceeds a few seconds" (Persinger & Cameron, 1986, p. 49). In some cases, phenomena such as bites, cuts, scratches, welts, and possession-like trances have also been documented (e.g., Amorim, 1990; Mulacz, 1999).

This depiction of "ghostly episodes" is cogent, because it reduces these anomalies to neutral, behavioral-like components and terms. Moreover, consistent with other suppositions (Belz & Fach, 2015; Dixon, 2016; Houran, 2002; Laythe & Owen, 2013), it maintains a clean distinction between *Subjective* (*S*, internal or psychological) and *Objective* (*O*, external or physical) incidents when assessing cases. It

1 Send correspondence to: James Houran, Ph. D., Laboratory for Statistics and Computation, Polytechnic Institute of Management and Technology, Rua Cabo Borges (a` Av. República) 4430-646, Vila Nova de Gaia, Portugal, jim_houran@yahoo.com

also emphasizes overlapping features between haunts and poltergeists (Williams & Ventola, 2011, pp. 14-15), thereby suggesting the possibility that both episodes share an underlying phenomenon or set of mechanisms (Evans, 1987, 2001; Houran, 2000; Hufford, 1982). However, a curious feature not captured by Persinger and Cameron's (1986) synopsis is that episodes also exhibit "focusing effects" – i.e., to varying degrees incidents simultaneously center around certain places or objects and people (Roll, 1977). Taken altogether, these patterns might argue for an *interactionist* phenomenon defined by an interplay of "the right people in the right environments" (e.g., Lange & Houran, 2001a; Laythe, Houran, & Ventola, 2018).

Self-report measures generally agree on the base experiences characterizing witness narratives (Houran et al., 2019), but ongoing research has been stifled by the lack of a specific and standard operationalization. Consequently, confirming the *S/O* factor structure of these occurrences (i.e., one or two dimensions representing subjective and/or objective events) and conducting meaningful cross-study comparisons on putative causes or correlates remain elusive. In this paper, we address this issue via Modern Test Theory analyses of purportedly authentic reports of a spontaneous and anomalous nature versus a set of control accounts (cf. Houran & Brugger, 2000), i.e., narratives derived from contexts of strong suggestion or expectation, as well as narratives by individuals instructed to fabricate accounts.

Readers unfamiliar with this class of analytics are referred to Houran's (2017, pp. 191-193) summary of features and benefits, since it is well-established (Bond & Fox, 2015; Wright & Masters, 1982; Wright & Stone, 1979) that self-report instruments developed with Classical Test Theory (CTT) are often tainted by serious measurement problems. Unfortunately, most instruments in the paranormal belief and anomalous experience literature fall in this category (e.g., Dixon, 2016; Irwin, Dagnall & Drinkwater, 2013; Jinks, 2012; Sen & Yesilyurt, 2014; Schofield, Baker, Staples, & Sheffield, 2018; Storm, Drinkwater, & Jinks, 2017; Tobacyk, 2004). By contrast, Lange, Irwin, and Houran (2000) introduced a series of psychometric analyses they described as a "top-down purification" process. This method is used increasingly in consciousness studies (e.g., Irwin, & Marks, 2013; Lange, 2017; Preti, Vellante, & Petretto, 2017), and it combines Rasch (1960/1980) scaling with the removal of age- or gender-related responses biases. Controlling for these is critical, because statistical theory (Stout, 1987) and computer simulations (Lange et al., 2000) alike demonstrate that response biases can lead to spurious factor structures of constructs, significant distortions in scores, and consequently erroneous reliability and validity findings.

In addition to improved quality control and model-building, Rasch scaling often produces critical insights into the constructs being considered. For instance, Lange's research programs (see e.g., Lange, 2017) have often found that differences in the phenomenology of biopsychosocial constructs, like depressive symptoms or expressions of romantic love, are partly rooted in respondents' age, gender, or cultural learning (e.g., Lange, Houran, & Li, 2015; Lange, Thalbourne, Houran, & Lester, 2002). Therefore, this paper applied Modern Test Theory to clarify the degree to which nuances in the phenomenology of ghostly episodes reflect idiosyncrasies of experients versus the nature of the construct(s) per se.

A Rasch Scaling Primer

Rasch scaling and related Item Response Theory models (e.g., van der Linden & Hambleton, 1997) provide valuable information that goes far beyond that provided by standard (i.e., raw-sum-based) numerical test scores. Instead, the individual items are seen as independent probabilistic sources of information that combine to assess a common underlying latent variable. Respondents' answers are explicitly modeled as the juxtaposition of items' "difficulty" and respondents' trait levels. This formulation has several advantages, as it allows the derivation of maximum-likelihood estimates of respondents' trait levels expressed at an interval-level of measurement. At the same time, Rasch scaling provides valuable quality control indices, including items' fit to the Rasch model and statistics that indicate whether items and test scores are systematically biased *for* or *against* subgroups of respondents. Although misfit and bias obviously impede measurement, recent research indicates that their quantification also creates powerful predictors of considerable diagnostic value (Lange & Houran, 2015; Lange, Martínez-Garido, & Ventura, 2017). This possibility is explored here as well.

More formally, Rasch scaling of binary items models a respondent's (*j*) affirmation of an item (*i*) as the probabilistic outcome of two factors: the respondent's trait level (*Tj*) and the level (*Di*) at which item *i* assesses the trait. *Di* is also called the "item difficulty" or its "location." The value of *D decreases* as items receive more affirmative answers. Given the preceding terminology, the following equation describes the Rasch model for binary items:

$$ln(Pij/(1-Pij)) = Tj - Di \quad (1)$$

where *Pij* denotes the probability that person *j* will respond affirmatively to item *i*. The parameters *T* and *D* in Equation 1 are expressed in a common unit (called "logits"), as is defined by the log-odds ratio on the left-hand side of Equation 1. Accordingly, logits denote the locations of items within the Rasch hierarchy, with higher values indicating higher positions (or greater difficulty) on the scale. The values of *T* and *D* can be derived from sample data using iterative maximum likelihood estimation procedure for which we used Linacre's (2018a, 2018b) *Winsteps* and *Facets* Rasch scaling software. Detailed introductions to Rasch scaling are readily available to interested readers (Bond & Fox, 2015; Lange, 2017; Wright & Mok, 2000).

Examples. Note that in Equation 1 whenever *Tj* equals *Di* (i.e., respondent *j*'s trait level equals the difficulty of item *i*), then *Pij = 0.5*. In other words, an item's difficulty *D* is the point where endorsement and non-endorsement are equally likely. Moreover, $Pij < 0.5$ whenever $Tj < Di$ and $Pij > 0.5$ whenever $Di < Tj$. Consider for instance, the statement "I had the mysterious feeling of being watched, or in the presence of an invisible being or force," which is an item of the SSE questionnaire to be discussed later (see Appendix B). If we assume that $D = 0.84$, then people with trait level $T = 0.84$ have a 50% chance of reporting feelings of being watched or sensing an invisible force or being. Those with lower trait levels ($L < 0.84$) are less likely to do so, i.e., their $p < 0.5$, but notice that they still might do so. Those with higher trait levels than 0.84 report such experience with greater likelihood (i.e., $p > 0.5$). But again, it is never certain that they will report the experience. In general, whenever *T* increases, so does *P*. Also, whenever some item *a* is "easier" than another item *b* (i.e., $Da < Db$) then $Pa > Pb$,

given *T*. In other words, for all respondents with trait level *T* easier items are always more likely to be endorsed than are harder ones.

Model Fit. By solving for *Pij* in Equation 1, it is possible to compute respondents' expected ratings and the standard deviation thereof (Wright & Masters, 1982) given *T* and *D*. The residual *Resij* is defined as the difference between the actual answers (coded as integer values 0 or 1) and the (average) expected answer – i.e., a real-valued number ranging from 0 to 1. That is,

Resij = Actual Answer – Expected Answer (2)

ZResij is the standardized form thereof (i.e., with *M* = 0 and *SD* = 1). The *Resij* values can be factor analyzed to detect the presence of secondary factors that threaten unidimensionlity. Further, items with high (absolute) residuals are said to show "misfit," i.e., they do not act according to Equation 1. If the *ZRes* follow a chi-square distribution then their squared sum follows a chi-square distribution with *df = N*, where *N* denotes the number of persons who took the item. The average *ZRes2* reflects the items' Outfit. Since this sum equals the degrees of freedom, dividing by *N* yields a statistic with an expected value of 1. Experience indicates that Outfit values ranging from 0.7 to 1.4 are generally acceptable (Linacre, 2018a), while larger values suggest misfit.

DIF and Res. Linacre's *Facets* (2018b) parameter estimation software produces statistical tests to check the equality of the items' difficulties across subgroups. Systematic variation in *D* across subgroups is called "differential item functioning" (*DIF*), or response bias. Such *DIF* is also captured by the observations' *Resij* (Equation 2), a negative *Resij* value implies that an answer is unexpectedly low, whereas a positive *Resij* implies the opposite. Each person's 32 *Resij* were computed via Linacre's (2018a) *Winsteps*, and they were added to respondents' data records.

This Study

We aimed to develop a "top-down purified" Rasch measure that quantifies the phenomenology of ghostly episodes while controlling for potential response biases related to age and gender. Ideally, this inventory would be suitable for surveys, fieldwork studies, and investigations that codify free-response data or spontaneous case material for quantitative analysis. In this way, researchers will have a standard method of operationalizing, analyzing, or scrutinizing accounts. Meeting these goals also informs model-building or theory-formation, since we tested four hypotheses:

- *Hypothesis 1:* The phenomenology of "spontaneous" ghostly episodes will show a predictable (cumulative) behavioral pattern, as evidenced by conformity to a unidimensional probabilistic Rasch model that subsumes *S/O* classes of events within a single construct (Houran & Lange, 2001, 2009; Houran, Wiseman, & Thalbourne, 2002).

- *Hypothesis 2:* The Rasch residuals of spontaneous accounts will demonstrate significant differences compared to four control groups (Lifestyle, Primed, Fantasy, and Illicit, defined in the Method section). This hypothesis took the form of statistical tests for group-related *DIF*.

- *Hypothesis 3:* It seems likely that the aforementioned *DIF* is sufficiently powerful to produce pronounced group-related distortions that serve as "signatures" of group membership. It should thus be possible to infer respondents' group membership (i.e., Spontaneous, Lifestyle, Primed, Fantasy, and Illicit) from the Rasch residuals of their responses.

- *Hypothesis 4:* Previous studies (Lange & Houran, 2015; Lange et al., 2017) indicate that Rasch residuals predict class membership better than raw-score observations. We expected to replicate this finding here.

Method

Respondents

Data derived from a convenience sample of 621 participants from the United States and Great Britain who completed an online survey. The overall sample (M_{age} = 40.89, *SD* = 12.49; r*ange* = 18 to 73) consisted of 459 women, 156 men, and 6 transgender participants. Accounting for our five conditions (see below), 330 women and 92 men, and 4 transgender participants with a M_{age} of 41.71 (*SD* = 12.15, r*ange* = 18 to 73) completed the Spontaneous condition. The Primed condition comprised 40 women and 10 men with a M_{age} of 41.75 (*SD* = 12.11, r*ange* = 18 to 66). The Lifestyle condition comprised 31 men and 14 women, with a M_{age} of 43.60 (*SD* = 11.48, r*ange* = 22 to 69). The Fantasy condition comprised 30 women and 20 men with a M_{age} of 34.98 (*SD* = 12.75, r*ange* = 18 to 59). Finally, the Illicit condition contained 38 women, 20 men, and 2 transgenders with a M_{age} of 37.18 (*SD* = 13.86, *range* = 18 to 66). This study was approved by the Human Research Ethics Committees at Indiana University.

Various tactics were unsuccessful in increasing and balancing sample sizes across the groups, so data collection ceased after it stagnated. Although we acknowledge the limitations of our samples, Wright and Douglas (1975) noted that pilot studies with as few as 30 observations are often useful, and this agrees with later sampling guidelines by other authorities (e.g., Kruyen, 2012; Linacre, 2002). In fact, small samples are often sufficient to identify inferior items from a Rasch perspective (Wright & Stone, 1979).

The Survey of Strange Events

Our previously collated set of haunt and poltergeist anomalies (Houran et al., 2019) was the basis for a new 32-item checklist that uses language accessible to a wide range of respondents. Readability statistics (via readable.io/text/) indicate that the questionnaire met a 10th grade level of comprehension (contact the authors for details; for an introduction to this topic see Kouamé, 2010).

The row entries of Table 1 list the items on the binary (True = 1/ False = 0) checklist, which we titled the *Survey of Strange Events* (SSE) to frame the inventory in a more neutral context (see Appendix B). Likewise, the wording of items was intended to describe the various events in mostly dispassionate terms, without prejudicial modifiers like "paranormal" that denote an etiology for the events. This approach follows other researchers who distinguish anomalous experiences from their interpretations (David, 2010; Irwin et al., 2013).

Table 1

Summary of Rasch Scaling Results for the SSE Items by Respondent Condition.

		Spontaneous Group Only				*Item Location Difference Relative to Spontaneous Group*				
Item No.	*Brief Description*	*Location*	*Outfit*	*Gender DIF*	*Age DIF*	*Primed*	*Lifestyle*	*Fantasy*	*Illicit*	χ^2
1	Non-descript visual form*.	-0.62	1.14	*0.63*	-0.05	0.28	0.02	0.39	0.68	7.44
2	Obvious apparition*.	-0.51	1.07	0.19	0.32	0.40	0.54	1.04	-0.34	15.64
3	Alive-looking apparition*.	-0.47	1.08	0.33	0.24	0.41	1.04	0.87	0.09	16.34
4	Pleasant odor*.	0.03	1.01	*-0.67*	*0.60*	0.76	0.41	0.63	1.28	23.20
5	Unpleasant odor*.	0.42	0.88	-0.07	-0.02	-0.07	0.04	-0.06	-0.25	0.83
6	Positive feeling*.	0.10	1.03	-0.02	0.43	0.47	0.61	0.62	2.30	*39.75*
7	Negative feeling*.	-0.60	1.17	-0.17	-0.21	-0.31	0.13	0.73	1.17	*23.98*
8	Odd bodily sensations*.	-0.47	1.13	0.24	*-0.63*	-0.16	-0.11	0.66	1.02	*18.4*
9	Mysterious taste*.	1.08	0.84	0.14	-0.06	0.27	1.29	0.48	0.99	11.81
10	Possessed by outside force*.	0.84	0.79	0.34	-0.33	0.16	-0.21	0.22	-0.80	9.07
11	Mystical-type beings*.	1.07	0.78	0.51	-0.06	-0.01	-0.10	0.35	-0.43	3.15
12	Folklore-type beings*.	1.61	0.71	-0.71	0.25	-0.38	0.08	0.12	-0.59	3.82
13	Communicated with dead/force*.	0.03	0.99	-0.19	0.10	-0.77	-1.58	0.81	0.30	*35.21*
14	Sensed presence*.	-1.59	*1.47*	0.04	-0.53	-0.09	0.42	-0.08	1.80	*42.93*
15	Déjà vu*.	-1.65	*1.48*	0.00	-0.34	-0.15	0.80	0.14	1.68	*39.04*
16	Recognizable sound.	-0.62	1.11	-0.12	-0.08	0.15	-0.20	-0.50	0.26	4.36
17	Non-descript sound.	-1.17	1.25	-0.49	0.09	-0.24	-0.21	-0.46	0.52	6.14
18	Recorded & recognizable sound.	0.24	0.86	0.58	0.19	-0.60	-1.79	-1.07	-1.52	*51.09*
19	Recorded & non-descript sound.	0.16	0.90	0.05	0.37	-0.59	-1.27	-0.50	-1.03	*25.61*
20	Cold area.	-0.80	1.17	0.09	-0.10	-0.35	0.11	-0.02	0.42	3.60
21	Hot area.	0.72	0.85	0.43	-0.06	-0.15	-0.20	0.47	0.06	2.18
22	Object teleportation.	-0.10	0.98	0.15	-0.29	0.71	1.04	-0.18	-0.17	11.52
23	Object movement.	0.05	0.90	-0.10	-0.04	0.62	-0.28	-1.11	-1.25	*28.39*
24	Object levitation.	0.65	0.78	0.48	0.31	0.77	0.28	-0.93	-1.38	*30.8*
25	Erratic functioning – electronics.	-0.62	1.09	-0.15	*0.72*	-0.04	-0.46	-0.58	-0.20	5.12

26	Recorded image.	-0.05	0.94	-0.45	0.33	-0.62	-0.45	-0.44	-1.06	*14.77*
27	Erratic functioning - plumbing.	0.90	0.83	-0.35	-0.21	-0.01	0.28	-0.76	0.50	9.95
28	Object breakage.	0.51	0.80	0.28	-0.20	0.63	1.08	-0.91	-1.61	*39.7*
29	Anomalous breeze.	-0.73	1.16	-0.19	-0.38	0.15	0.44	-0.29	0.69	9.13
30	Anomalous fires.	1.71	0.59	-0.44	0.29	-0.68	0.91	-0.26	-2.27	*50.08*
31	Non-threatening touch.	-0.55	1.12	-0.31	-0.14	-0.04	-0.19	0.35	0.63	6.95
32	Threatening touch.	0.44	0.83	0.20	-0.33	-0.29	-0.29	-0.08	-1.56	*23.92*

Respondent Sub-Groups

We created five respondent groups to complete the SSE, along with a demographic Face Sheet. An algorithm facilitated the automated assignment of respondents into these groups:

Group 1: **Spontaneous** – respondents from the general population who reported an apparently sincere and unprimed episode, whereby they were not actively seeking paranormal events. Such accounts can arguably be judged as more likely to be anomalous and thus of potential evidentiary value to parapsychology, i.e., addressing Schmeidler's (2001) question, "...has the skeptic discussed only the easier cases and neglected the hard ones?" (p. 308) — a sentiment strongly echoed by Stokes (2017a, 2017b). The instruction set was: "*Please recall vividly a time when you visited or lived in a specific place without seeking or wanting anything paranormal to happen. But there, you experienced strange or unexplained events that some people would call a ghost or haunting...*"

Group 2: **Primed** – respondents from the general population who had anomalous experiences during commercial ghost tours, which are thus likely attributable to expectation or suggestion (French, Hague, Bunton-Stasyshyn, & Davis, 2009; Lange & Houran, 1997) or clear-cut demand characteristics (Orne, 1962; Slosson, 1899). The instruction set was: "*Please recall vividly a recent commercial ghost tour or commercial public ghost event where you visited a specific location and experienced strange or unexplained events that some people would call a ghost or haunting...*"

Group 3: **Lifestyle** – respondents with active memberships in self-styled ghost-hunting or ghost-tour groups (so-called "hauntrepreneurs," e.g., Hill, 2017; Potts, 2004) who are likely under the influence of strong context effects like pervasive paranormal belief or demand characteristics (French, 1992; Harte, 2000; Houran, 2000). The instruction set was: "*Please recall vividly a recent investigation or tour to a specific location where you experienced strange or unexplained events that convinced you that ghosts were real...*"

Group 4: **Fantasy** – respondents with no prior ghostly experiences who were asked to imagine what a vivid and personal experience would be like, thus eliciting narratives that are likely to be intuitively-generated (cf. Sinclair & Ashkanasy, 2005), i.e., creatively constructed partly from tacit knowledge

accumulated through experience and cultural learnings, combined with a psychophysiological ability to access sensory and affective elements (MacKinnon, 1971; Rugg, 1963; Zausner, 1988). The instruction set was: "*Please imagine vividly what it would be like to live in a genuine 'haunted house' by yourself, or with others, for a period of one month. Try to visualize in great detail the kind of strange or unexplained events you might experience...*"

Group 5: **Illicit** – respondents with no prior ghostly experiences asked to concoct a bogus but seemingly convincing account. This slightly resembles the Fantasy group above, except that narratives here would arguably cater more to social approval or cultural norms, especially as related to how paranormal themes are characterized in popular culture (Booker, 2009; Edwards, 2001; Goldstein, Grider, & Thomas, 2007). The instruction set was: "*Please imagine vividly the following scenario—You are answering an ad for a new paranormal TV show. The producers are seeking only sincere witnesses to interview on camera about their experiences living in a genuine 'haunted house.' You do not actually live in a haunted house but there may be a number of different reasons why you might want to appear on the TV show. Respond to the survey below in a way that you think would best convince the producers that you really do live in a genuine haunted house and thereby win a spot on the show...*"

Group Classifications

We addressed the classifications required by Hypotheses 3 and 4 using the Python-based suite of Machine Learning models included in *Sci-Kit Learn* (Pedregosa et al., 2011). Given their wide usage, we included Logistic Regression generalized to multiple categories, Support Vector Machines, Decision Trees, and Linear Discriminant Analysis (LDA). For background information, see Hastie, Tibshirani, and Friedman (2009).

Using the standard machine learning approach, the 621 cases were randomly divided into a Training and Validation set of respondents, with approximately 33 and 67% of the cases, respectively. Next, the classification models mentioned above were fitted optimally to the Training set. The resulting models were then applied to the validation set without any parameter changes. Since the models typically capitalize on irrelevant peculiarities of the Training set, resulting in over-fitting, we will report *only* the accuracy of prediction in the Validation set.

Results

Hypothesis 1

Table 1 summarizes the Rasch analyses results for the SSE's 32 items, calibrated on data from the intended population only (i.e., the Spontaneous group). Supporting Hypothesis 1, and conceptually replicating previous findings, data showed acceptable Outfit, as just two items yielded fit values that marginally exceeded the criterion value of 1.4 (i.e., 1.47 and 1.48, for Items 17 and 18, see first numerical column). Also, factor analysis of the items' Rasch residuals supported a unidimensional model, since the first (and most important) residual factor accounted for less than 5% of the total variance. Next, tests for item shifts (or differential item functioning: DIF) across age and gender showed no statistically significant effects (all $p > 0.10$) and the magnitudes of the age and gender items shifts are shown labeled DIF. These

findings indicate that the item hierarchies for "women vs men" and "older vs. younger" (defined by a median split) percipients do not differ appreciably. Thus, for Spontaneous respondents, the SSE apparently defines a stable, probabilistic hierarchy that varies little across experients' age or gender.

In Table 1 *S* type items are marked with an asterisk, but not the *O* items. Comparison of the *S* and *O* items reveals interesting patterns for model-building and theory-formation. In particular, not only do the *S/O* items constitute a single factor, but the two putative experience types were interspersed in the Rasch hierarchy as opposed to being markedly or consistently disconnected from one another. Therefore, it does not seem that the phenomenology of ghostly episodes begins, evolves, or ends with experiences that are exclusively psychological or physical in nature. Instead, there is a reasonably balanced mix of these events across the Rasch hierarchy.

That said, events in the SSE hierarchy can be divided into ~1-logit increments yielding three sets of experiences that are comparatively *common* (-1.6 to -.5), *less common* (-.5 to .5), and *rare* (.5 to 1.7). Considered this way, and rough-coding the contents of the SSE items, we find that Auditory and Sensation (internal feelings) experiences figure heavily in common events but become absent in low-probable events. Visual-related anomalies become more probable in an episode as event rarity increases. Moreover, as might be expected, commonly-endorsed items seem to be readily explainable, since they refer to few, if any, events that might be interpreted by experients as clear and convincing evidence for paranormality — or, at least, meet a threshold for obvious *aberrant salience* (Irwin, 2014; Irwin, Schofield, & Baker, 2014). Conversely, as we move to the less-probable categories, events increasingly are more paranormal-looking, thus requiring more complex explanations.

Last, an even closer look at the items' distribution exposes apparent "clustering" effects whereby some SSE items have logit values (i.e., locations in the Rasch hierarchy) within close range of each other (±.05 logits). That is, some specific events tend to coincide closely with certain other events. We identified six such clusters in the SSE hierarchy — half of these clusters show *S/O* entanglement, although *S* events dominate virtually all the clusters. Additionally, *S* events are almost always primary in these clusters. Such clustering effects occur across the Rasch hierarchy, and we explore possible explanations in the Discussion.

Appendix A shows the conversion of raw-sum scores to Rasch-scaled scores, which quantifies the *intensity* of ghostly episodes at an interval-level (Rasch person reliability = 0.87). For convenience, the original Logit scale was transformed to yield a scale score with a mean 50 and *SD* = 10, as based on the data in the Spontaneous group. On this scale men and women differed little (M_{men} = 50.6 vs. M_{women} = 49.4), as did younger and older individuals ($M_{younger}$ = 49.7 vs. M_{older} = 50.3), and neither difference is statistically significant ($p > 0.10$) as computed via Linacre's (2018b) software. However, the five respondent groups ($M_{Spontaneous}$ = 51.7, M_{Primed} = 52.3, $M_{Lifestyle}$ = 50.6, $M_{Fantasy}$ = 49.43, $M_{Illicit}$ = 45.9; $\chi^2(df = 4, N = 622) = 330.55$, $p < .01$) showed significant mean differences.

Hypothesis 2

We found strong support for Hypothesis 2. First, an overall test for item shifts across the five groups revealed statistically significant differences in their item hierarchies (χ^2 (df = 160, N = 622) = 613.87, $p <$ .001). Table 1 shows the extent to which the SSE items' locations in the Primed, Lifestyle, Fantasy, and

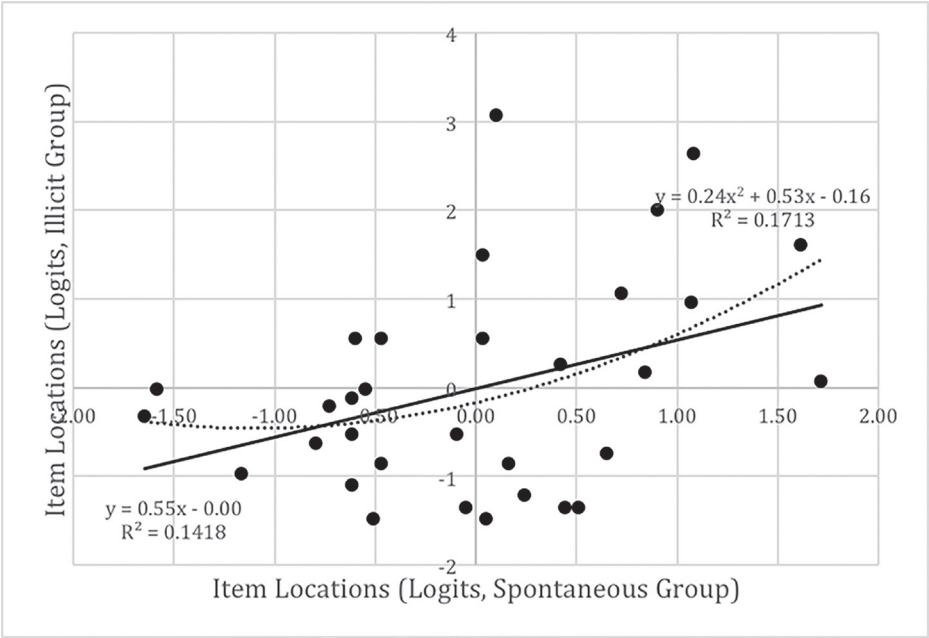
Illicit groups differed from the locations obtained in the Spontaneous group (i.e., relative to the item location shown as the first numerical column). Positive values indicate that the specific event was *under*-reported, whereas negative values indicate that the event was *over*-reported.

Second, as is specified by the underlined values in the χ^2 column of Table 1, further tests revealed that the locations of 17 of the 32 items differed significantly across the four groups ($p < .01$). For instance, Item 16 ("I communicated with the dead or other outside force") shows disparate trends across the control groups, i.e., compared to the Spontaneous group, it is a much rarer experience in "haunt-seeking" contexts (Primed and Lifestyle groups). Interestingly, however, it was specified by those in the Fantasy and Illicit conditions. In contrast, Item 7 ("I heard on an audio recorder mysterious sounds that could be recognized or identified, such as ghostly voices or music, with or without singing") showed all negative values across the control groups, indicating that endorsement of this anomaly is "harder" (or rarer) in the Spontaneous group.

To investigate the possible existence of meta-patterns, Figure 1 plots the item location obtained in the Illicit group (Y-axis) against the locations found in the Spontaneous group (X-axis). Recall that items' locations *Di* indicate the point where their probability of endorsement is 50%. Clearly, there is no simple relation between these two item sets ($r = 0.37$) and adding a quadratic component had little effect ($r = 0.41$). Instead, it appears that the nature of the relation changes with greater item difficulty in the Spontaneous group. Specifically, for X < 0 the locations in the Illicit group are essentially constant, with a mean around -0.47 logits in small range. However, for X > 0 we get a "buckshot" pattern over a much wider Y-range.

Specifically, Illicit respondents checked "frequently-reported" (or "easy") experiences ($N = 15$) with very similar frequency. This suggests that Illicit respondents regarded all the "easy" events as similarly plausible, i.e., low variation ($SD = 0.60$ logits), and their locations show a small negative correlation with those obtained in the Spontaneous group ($r = -.20$). By contrast, Illicit respondents express definite

Figure 1. Item Locations in the Spontaneous Condition (X-axis) vs. Illicit Condition (Y-axis)



ideas about the plausibility of events defined by the 17 "harder" items (X > 0), i.e., they assessed some items as very plausible (Y < 0), while others were deemed quite implausible (Y > 0), and this yields far greater variation (*SD* = 1.41 logits). Again, however, there is little correlation between items' locations in the two groups (*r* = .26).

We note that the group difference in the *SD*s of the "easier vs. harder" item locations reaches statistical significance, $F(14,16) = 5.52$, $p < .001$. Further, the 32 SSE items' locations vary more in the Illicit group (*SD* = 1.18) than in the Spontaneous group (*SD* = 0.82, $F(31,31) = 2.10$, $p < .05$). Thus, those trying to fabricate accounts of ghostly episodes fail to achieve the proper difficulty-gradations of the experiences listed in the SSE. In particular, they either *over*- or *under*-estimate the incidence of events in a seemingly random pattern, especially when dealing with rare events.

Hypotheses 3 and 4

It is reasonable to ask whether the item-shifts across the different conditions outlined above are powerful and systematic enough to predict respondents' group membership. In doing so, the use of residuals was compared to using raw scores. In both cases, age and gender were included as predictors as well. Recall that we used a variety of classification models (e.g., LDA, Logistic Regression, Support Vector Machines, and Decision Trees), and the main interest is models' predictive quality when applied to the validation set (see Table 2).

Consistent with Hypothesis 3, all predictive approaches performed well, although LDA slightly outperformed the other approaches (see Table 2). The bottom row ("Accuracy") of Table 3 indicates that whereas the overall LDA accuracy was quite high (72%), performance was not uniform across groups, the Primed, Lifestyle, and Fantasy groups performed very poorly. In fact, when combined, only 13 % of the cases in these three groups were correct, i.e., the classification methods distinguished mostly between the Spontaneous vs. Illicit groups. As in earlier research (Lange & Houran, 2015; Lange et al., 2017), Table 2 showed that residuals were consistently better predictors than raw scores, supporting Hypothesis 4 — albeit by very small margins.

Table 2
Proportions of Correct Predictions by Method and Predictor Variables (Validation Group Only)

	Predictor Variables	
Method	*Residuals*	*Raw*
Linear Discriminant	0.72	0.70
Logistic Regression	0.71	0.69
Support Vector Machine	0.70	0.67
Decision Tree	0.68	0.67

The preceding strongly suggests that some groups could be combined to improve performance. In a first attempt, we contrasted the Spontaneous group against all four others combined, and this yielded a minor increase in performance (78%) based on the best predictor (LDA). Yet, when contrasting the Spontaneous group against all others, classification correctness rose to 94% using logistic regression.

Interestingly, in the latter case the Decision Tree approach performed very well too (accuracy = 91%). Figure 2 shows a simplified form of this tree that describes an interesting heuristic. First, individuals who do not report item #30 ("Fires have started mysteriously") have an 87% chance of belonging to

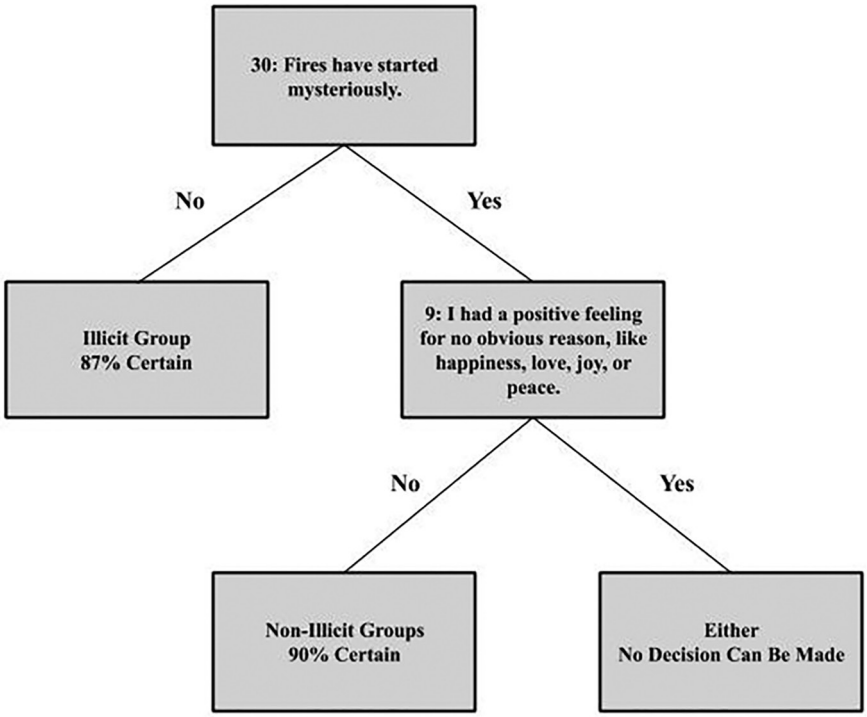
Table 3
Counts in Confusion Table for Prediction Based on Linear Discriminant Analyses (Validation Set only)

Actual	Predicted Group					
Group	*Spontaneous*	*Primed*	*Lifestyle*	*Fantasy*	*Illicit*	Total
Spontaneous	125	0	1	1	3	130
Primed	13	0	2	0	2	17
Lifestyle	14	0	2	2	2	20
Fantasy	10	0	0	5	1	16
Illicit	6	0	0	1	16	23
Total	168	0	5	9	24	206
Accuracy per Column (%)	74	0	40	56	67	

the Illicit group. Second, those who do endorse item 30 but not item #9 ("I had a *positive* feeling for no obvious reason, like happiness, love, joy, or peace") with 90% certainty belong to one of the non-Illicit groups (i.e., the Spontaneous, Primed, Lifestyle, or Fantasy groups combined). Note that the "Either" leaf (i.e., membership cannot reliably be decided) contains just 4% of the cases.

We also note that overall tests to detect shifts in items' Rasch D parameters across groups showed highly significant Respondent Group × Item interactions. Specifically, the Group × Gender, $\chi^2(df = 320, N = 622) = 747.6$); Group × Age, $\chi^2(df = 320, N = 622) = 746.4$; and Group × Gender × Age, $\chi^2(df = 640, N = 622) = 984.10$ interactions with the items' locations was significant at $p < .001$. Inclusion of these interactions in the computation of the residuals (see e.g., Lange et al., 2017) would likely improve the accuracy of predicting respondents' group membership. However, we judged our samples in all but the Spontaneous group as too small for robust results in this respect, especially since separate Training and Validation groups are required to fit the augmented predictive models. Instead, we propose to pursue this issue in future research.

Figure 2. Simplified Decision Tree to Differentiate Spontaneous and Illicit Accounts



Discussion

Despite the fleeting nature of ghostly episodes (Persinger & Cameron, 1986), our findings suggest that their phenomenology can be reliably quantified akin to other psychometric variables in the social and biomedical sciences and accordingly is amenable to rigorous statistical scrutiny. Moreover, we substantiated a new facet of study — i.e., beyond the *frequency* of ghostly episodes and the *variety* of their properties, there is now a standardized (internal-level and bias-free) measure of their *intensity*. Consequently, we can draw several conclusions about the phenomenology of the accounts in our sample, as operationalized and modeled here:

- They exhibit a predictable behavior pattern, consisting of a unidimensional and probabilistic (cumulative) hierarchy of core or base events.
- This hierarchy subsumes Psychological Experiences and Physical Manifestations within a common dimension. Therefore, the often presumed two-factor model using *S/O* classifications is not supported.
- The Rasch residuals of responses on the SSE are sufficiently powerful to demarcate the phenomenology of "Spontaneous" accounts from a set of "impostor" (or control) accounts. As in earlier research (Lange et al., 2017; Lange & Houran, 2015), the prediction of respondents' group membership was slightly more successful when using items' residuals versus raw scores. Although

a very simple decision tree approach proved quite successful in identifying impostor cases, the finding is specific to the present context and additional study is needed.

We should stress the provocative implications of ghostly episodes as a *unidimensional* construct. In contradiction to prior thinking (e.g., Dixon, 2016; Dixon, Storm, & Houran, 2018; Houran, 2002; Houran, Wiseman, & Thalbourne, 2002), this finding suggests there is neither a simple nor straightforward distinction between *S/O* categories. It is unclear what this single factor ultimately represents, but several hypotheses come to mind: (i) the two seemingly different classes of events share a common etiology, either entirely psychological or physical in origin, and thereby representing qualitatively different manifestations as the intensity of the core construct increases – not unlike sneezing, sore throat, nasal congestion, and fever as connected and worsening symptoms of a cold virus; (ii) the two classes of events are connected by perceptual abilities or attentional biases (e.g., Lange & Houran, 2001a; Laythe, Houran, & Ventola, 2018); (iii) the two classes of events are connected by narrative or interpretive processes (e.g., Baker & Bader, 2014; Eaton, 2019); or (iv) some combination of these or other variables.

Irrespective of the sources(s) for the core events, we contend that our Rasch analyses corroborate previous conceptual work (Nisbet, 1979; Palmer, 1974, Playfair, 1980; Pratt & Palmer, 1976) and empirical research (Houran & Lange, 2001; Houran, Wiseman, & Thalbourne, 2002) in modeling ghostly episodes as a type of *syndrome*, i.e., a set of signs or symptoms that occur together to characterize a particular abnormality or condition (British Medical Association, 2018). This idea speaks to a wealth of literature on the relation of paranormal ideations to psychiatric illness and symptom perception, which many authorities deem an important area of exploration (Bentall, 2000; Houran, Kumar, Thalbourne, & Lavertue, 2002; Jawer, 2006, Jawer & Micozzi, 2009; Mathijsen, 2016; Neppe, 1992; Schofield & Claridge, 2007). A syndrome framework is also broadly consistent with Lange and Houran's (1998, 1999, 2001a) hypothesis that biopsychosocial processes in haunts and poltergeist disturbances parallel those operating in cases of mass hysteria or contagious psychogenic illness (e.g., Chen, Yen, Lin, & Yang, 2003; Colligan, Pennebaker, & Murphy, 1982; Wessely, 1987, 2000).

On this latter point, the SSE items that exhibit clustering effects (i.e., within very close range in the Rasch hierarchy) might be construed as "flurries" of reported perceptions, which Jones and Jones (1994) noted can be a good index of behavioral contagion. To clarify, *contagion* is the triggering of successive perceptions due to priming and can occur with individuals (Houran & Lange, 1996) or groups (O'Keeffe & Parsons, 2010). Therefore, these clusters seem consistent with the premise that ghostly episodes involve "contagious" mechanisms (Houran & Lange, 1996; Lange & Houran, 2001a, 2001b; Nisbet, 1979). Alternatively, these clusters might reflect "syncretic perceptions" (Werner, 1934/1978, 1948; cf. Laythe et al., 2018), i.e., the dedifferentiation (or fusion) of perceptual qualities in subjective experience, e.g., *eidetic imagery* (fusion of imagery and perception, i.e., structural eidetic imagery); *physiognomic perception* (fusion of perception and feeling); and *synesthesia* (fusion of sensory modalities). The full import of these patterns will be unclear until they are proven to be robust via replication and explored with more detailed designs. This effort should include in-depth psychometric studies to elucidate whether individual differences in variables that can affect stimulus detection, interpretation, or concomitant response also help shape the phenomenology of ghostly episodes (e.g., Houran, Wiseman, & Thalbourne, 2002;

Langston & Hubbard, 2019; Laythe et al., 2018; Parra, 2018; Parra & Argibay, 2016). We are examining these ideas and will report the results in another paper.

These findings and ideas exemplify the type of research and model-building possible with the SSE. Of course, our conclusions are tempered by several limitations. First, the data were self-reported, and the underlying assumptions of the five respondent groups were not independently-corroborated. Second, our findings are based on responses specific to the SSE's wording. We neither claim that our articulations of the items were optimal, nor that they will generalize to cross-cultural contexts. Finally, results with convenience samples do not necessarily reflect those from large-scale, representative surveys of the general population. Moreover, we do not know whether item-shifts in the Rasch scale of the spontaneous group might be introduced by potentially important variables not considered here, such as social desirability bias (impression management), the latency between experiences and their documentation, different perceptual-personality profiles of witnesses, or variances across physical or social settings. Likewise, the results derived from those reporting single instances of anomalous experiences rather than multiple experiences over time or collectively across different witnesses within the same case. Obviously, the SSE hierarchy might alter in these scenarios. A separate report will address these types of issues, which arguably speak directly to the notion of ghostly episodes as an interactionist phenomenon.

Clearly much work lies ahead in establishing different validities and applications of the SSE. We suggest that it offers a standardized method for codifying and quantifying free-response narratives from historical or modern spontaneous cases, or similar data from controlled or quasi-experimental designs (e.g., Dixon, 2016; French et al., 2009; Houran, Wiseman, & Houran, 2002; Laythe, Laythe, & Woodward, 2017; Wiseman, Watt, Greening, Stevens, & O'Keeffe, 2002). We plan to pursue these and other avenues, although we also encourage independent studies and especially collaborative efforts. As such, the SSE is freely available for all researchers to use (with attribution). We hope this unprecedented psychometric approach opens new frontiers of innovative research designs and hypothesis-testing with which to explore our haunted brains and houses.

References

Amorim, M-A. (1990). The Guarulhos poltergeist: A reassessment of Andrade's (1984) monograph. *Journal of the Society for Psychical Research*, *56*, 193-207.

Baker, J. O., & Bader, C. D. (2014). A social anthropology of ghosts in twenty-first-century America. *Social Compass*, *61*, 569–593. doi.org/10.1177/0037768614547337

Belz, M., & Fach, W. (2015). Exceptional experiences (ExE) in clinical psychology. In E. Cardeña, J. Palmer, & D. Marcussion-Clavertz (Eds.), *Parapsychology: A handbook for the 21st century* (pp. 364-379). Jefferson, NC: McFarland & Co.

Bentall, R. P. (2000). Research into psychotic symptoms: Are there implications for parapsychologists? *European Journal of Parapsychology*, *15*, 79-88.

Booker, M. K. (2009). *Red, white and spooked: The supernatural in American culture*. Westport, CT: Praeger.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed). Mahwah, NJ: Lawrence Erlbaum. doi.org/10.4324/9781315814698

British Medical Association (2018). *BMA illustrated medical dictionary* (4th ed). London, England: DK Publishing.

Chen C. S, Yen, C. F, Lin, H. F., & Yang, P. (2003). Mass hysteria and perceptions of the supernatural among adolescent girl students in Taiwan. *Journal of Nervous and Mental Disease*, *191*, 122-123. doi.org/10.1097/01.NMD.0000050941.22170.A5

Colligan, M. J., Pennebaker, J. W., & Murphy, L. R. (Eds.) (1982). *Mass psychogenic illness: A social and psychological analysis*. Hillsdale, NJ: Erlbaum.

David, A. S. (2010). Why we need more debate on whether psychotic symptoms lie on a continuum with normality. *Psychological Medicine*, *40*, 1935-1942. doi.org/10.1017/S0033291710000188

Dixon, J. (2016). A case of ostensible poltergeist phenomena resulting in lingering haunt phenomena. *Australian Journal of Parapsychology*, *16*, 7-39.

Dixon, J., Storm, L., & Houran, J. (2018). Exploring ostensible poltergeist vs. haunt phenomena via a reassessment of spontaneous case data. *Australian Journal of Parapsychology*, *18*, 7-22.

Eaton, M. A. (2019). Manifesting spirits: Paranormal investigation and the narrative development of a haunting. *Journal of Contemporary Ethnography*, *48*, 155-182. doi.org/10.1177/0891241618756162

Edwards, E. D. (2001). A house that tries to be haunted: Ghostly narratives in popular film and television. In J. Houran & R. Lange (Eds.), *Hauntings and poltergeists: Multidisciplinary perspectives* (pp. 82-120). Jefferson, NC: McFarland & Co.

Evans, H. (1987). *Gods, spirits, cosmic guardians: A comparative study of the encounter experience*. Wellingborough, Northamptonshire, England: Aquarian.

Evans, H. (2001). The ghost experience in a wider context. In J. Houran & R. Lange (Eds.), *Hauntings and poltergeists: Multidisciplinary perspectives* (pp. 41-61). Jefferson, NC: McFarland & Co.

French, C. C. (1992). Factors underlying belief in the paranormal: Do sheep and goats think differently? *The Psychologist, 5*, 295-299.

French, C. C., Hague, U. Bunton-Stasyshyn, R., & Davis, R. (2009). The "haunt" project: An attempt to build a "haunted" room by manipulating complex electromagnetic fields and infrasound. *Cortex*, *45*, 619-629. doi.org/10.1016/j.cortex.2007.10.011

Goldstein, D. E., Grider, S. A., & Thomas, J. B. (2007). *Haunting experiences: Ghosts in contemporary folklore*. Logan, UT: Utah State University Press. doi.org/10.2307/j.ctt4cgmqg

Harte, T. M. (2000). Contextual mediation of perceptions during hauntings and poltergeist-like experiences: A replication and extension. *Perceptual and Motor Skills*, *91*, 451-459. doi.org/10.2466/pms.2000.91.2.451

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Modern applied statistics: Elements of statistical learning* (2nd ed.). New York: Springer. doi.org/10.1007/978-0-387-84858-7

Hill, S. (2017). *Scientifical Americans: The culture of amateur paranormal researchers*. Jefferson, NC: McFarland & Co.

Houran, J. (2000). Toward a psychology of 'entity encounter experiences.' *Journal of the Society for Psychical Research, 64,* 141-158.

Houran, J. (2002). Analysis of haunt experiences at a historical Illinois landmark. *Australian Journal of Parapsychology*, *2*, 97-124.

Houran, J. (2017). "Sheet happens!" Advancing ghost studies in the analytics age. *Australian Journal of Parapsychology*, *18*, 187-206.

Houran, J., & Brugger, P. (2000). The need for independent control sites: A methodological suggestion with special reference to haunting and poltergeist field research. *European Journal of Parapsychology*, *15*, 30-45.

Houran, J., Kumar, V. K., Thalbourne, M. A., & Lavertue, N. E. (2002). Haunted by somatic tendencies: Spirit infestation as psychogenic illness. *Mental Health, Religion & Culture*, *5*, 119-133. doi.org/10.1080/13674670210141061

Houran, J. & Lange, R. (1996). Diary of events in a thoroughly unhaunted house. *Perceptual and Motor Skills*, *83*, 499-502. doi.org/10.2466/pms.1996.83.2.499

Houran, J., & Lange, R. (2001). A Rasch hierarchy of haunt and poltergeist experiences. *Journal of Parapsychology*, *65*, 41-58.

Houran, J., & Lange, R. (2009). Searching for an optimal level of transliminality in relation to putative psi. *Journal of the Society for Psychical Research*, *73*, 92-102.

Houran, J., Laythe, B., O'Keeffe, C., Dagnall, N., Drinkwater, K., & Lange, R. (2019). Quantifying the phenomenology of ghostly episodes: Part 1 – Need for a standard operationalization. *Journal of Parapsychology*, *83*, 25-46. doi.org/10.30891/jopar.2019.01.03

Houran, J., Wiseman, R., & Thalbourne, M. A. (2002). Perceptual-personality characteristics associated with naturalistic haunt experiences. *European Journal of Parapsychology*, *17*, 17-44.

Hufford, D. J. (1982). *The terror that comes in the night: An experience-centered study of supernatural assault traditions*. Philadelphia, PA: University of Pennsylvania Press.

Irwin, H. J. (2014). Aberrant salience and motivation as factors in the formation of belief in scientifically unaccepted phenomena. *Journal of Scientific Exploration*, *28*, 633-645.

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. *Journal of Parapsychology, 77*, 39-53. doi.org/10.1037/t31377-000

Irwin, H. J., & Marks, A.D.J. (2013). The 'Survey of Scientifically Unaccepted Beliefs': A new measure of paranormal and related beliefs. *Australian Journal of Parapsychology*, *13*, 133-167.

Irwin, H. J., Schofield, M. B., & Baker, I. S. (2014). Dissociative tendencies, sensory-processing sensitivity, and aberrant salience as predictors of anomalous experiences and paranormal attributions. *Journal of the Society for Psychical Research, 78*, 193-206.

Jawer, M. (2006). Environmental sensitivity: Inquiry into a possible link with apparitional experience. *Journal of the Society for Psychical Research, 70*, 25-47.

Jawer, M. A., & Micozzi, M. S. (2009). *The spiritual anatomy of emotion: how feelings link the brain, the body, and the sixth sense.* Rochester, VT: Park St. Press.

Jinks, T. L. (2012). Paranormal and alternative health beliefs as quasi-beliefs: Implications for item content in paranormal belief questionnaires. *Australian Journal of Parapsychology*, *12*, 127-158.

Jones, M. B., & Jones, D.R. (1994). Testing for behavioral contagion in a case-control design. *Journal of Psychiatric Research*, *28*, 35-55. doi.org/10.1016/0022-3956(94)90035-3

Kouamé, J. B. (2010). Using readability tests to improve the accuracy of evaluation documents intended for low-literate participants. *Journal of MultiDisciplinary Evaluation, 6,* 132-139.

Kruyen, P. (2012) Using short tests and questionnaires for making decisions about individuals: When is short too short? pure.uvt.nl/portal/en/publications/using-short-tests-and-questionnaires-for-making-decisions-about-individuals(655f0d13-c333-4faa-aed2-b181f9fce71e).html, retrieved Jan 23, 2018.

Lange, R. (2017). Rasch scaling and cumulative theory-building in consciousness research. *Psychology of Consciousness: Theory, Research and Practice*, *4*, 135-160. doi.org/10.1037/cns0000118

Lange, R., & Houran, J. (1997). Context-induced paranormal experiences: Support for Houran and Lange's model of haunting phenomena. *Perceptual and Motor Skills*, *84*, 1455-1458. doi.org/10.2466/pms.1997.84.3c.1455

Lange, R., & Houran, J. (1998). Delusions of the paranormal: A haunting question of perception. *Journal of Nervous and Mental Disease*, *186*, 637-645. doi.org/10.1097/00005053-199810000-00008

Lange, R., & Houran, J. (1999). The role of fear in delusions of the paranormal. *Journal of Nervous and Mental Disease*, *187*, 159-166. doi.org/10.1097/00005053-199903000-00005

Lange, R., & Houran, J. (2001a). Ambiguous stimuli brought to life: The psychological dynamics of hauntings and poltergeists. In J. Houran and R. Lange (Eds.), *Hauntings and poltergeists: Multidisciplinary perspectives* (pp. 280-306). Jefferson, NC: McFarland & Co.

Lange, R., & Houran, J. (2001b). *Power laws and autoregressive catastrophes: The dynamic properties of poltergeist episodes*. Technical report to the Institut für Grenzgebiete der Psychologie und Psychohygiene (IGPP), Freiburg i. Br., Germany.

Lange, R., & Houran, J. (2015). "A picture is worth a thousand words:" Perceptual-personality profiling via a free-response imagery task. *North American Journal of Psychology, 17*, 387-402.

Lange, R., Houran, J., & Greyson (2015). Using computational linguistics to understand near-death experiences: Concurrent validity for the Near Death Experience Scale. *Psychology of Consciousness: Theory, Research, and Practice, 21, 79-89.* doi.org/10.1037/cns0000040

Lange, R., Houran, J., & Li, S. (2015). Dyadic relationship values in Chinese online daters: Love American style? *Sexuality & Culture*, *19*, 190-215. doi.org/10.1007/s12119-014-9255-0

Lange, R., Irwin, H. J., & Houran, J. (2000). Top-down purification of Tobacyk's Revised Paranormal Belief Scale. *Personality and Individual Differences*, *29*, 131-156. doi.org/10.1016/S0191-8869(99)00183-X

Lange, R., Martínez-Garrido, C., & Ventura, A. (2017). I'm scared to go to School! Capturing the effects of chronic daily fears on students' concept of self. *Journal of Applied Measurement. 18*, 420-433.

Lange, R., Thalbourne, M. A., Houran, J., & Lester, D. (2002). Depressive response sets due to gender and culture-based differential item functioning. *Personality and Individual Differences*, *33*, 937-954. doi:10.1016/S0191-8869(01)00203-3

Langston, W., & Hubbard, T. (2019). Shadow walking: Will a ghost walk tour affect belief in ghosts? *Journal of Parapsychology*, *83*, 47-68. doi.org/10.30891/jopar.2019.01.04

Laythe, B., Houran, J., & Ventola. (2018). A split-sample psychometric study of 'haunters.' *Journal of the Society for Psychical Research*, *81*, 193-218.

Laythe, B., & Owen, K. (2012). Paranormal belief and the strange case of haunt experiences: Evidence of a neglected population. *Journal of Parapsychology*, *76*, 79-108.

Laythe, B., & Owen, K. (2013). A critical test of the EMF-paranormal phenomena theory: Evidence from a haunted site without electricity-generating fields. *Journal of Parapsychology, 77*, 212-236.

Laythe, B. R., Laythe, E. C., & Woodward, L. (2017). A test of an occult themed séance: Examining anomalous events, psychosomatic symptoms, transliminality, and electromagnetic fields. *Journal of Scientific Exploration, 31*, 572-624.

Linacre, J. M. (2002), Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, *3*, 85-106.

Linacre, J. M. (2018a). *Winsteps*® [Computer Software]. Beaverton, OR: Winsteps.com.

Linacre, J. M. (2018b). *Facets*® [Computer Soft- ware]. Beaverton, OR: Winsteps.com.

MacKinnon, D. W. (1971). Creativity and transliminal experience. *Journal of Creative Behavior*, *5*, 227-241. doi.org/10.1002/j.2162-6057.1971.tb00893.x

Mathijsen, F. P. (2016). Anomalous experiences and schizotypy: Which comes first? The Hermit Crab syndrome hypothesis. *Mental Health, Religion & Culture*, *19*, 113-123. doi.org/10.1080/13674676.2015.1004627

Mulacz, P. (1999). Eleonore Zugun: The re-evaluation of a historic RSPK case. *Journal of Parapsychology*, *63*, 15-45.

Neppe, V. M. (1992). Anomalous experience and psychopathology. In B. Shapin & L. Coly (Eds.), *Spontaneous psi, depth psychology and parapsychology* (pp. 163-180). New York: Parapsychology Foundation, Inc.

Nisbet, B. C. (1979). A West Croyden "poltergeist:" Report and comments. *Journal of the Society for Psychical Research*, *50*, 229-237.

O'Keefe, C., & Parsons, S., (2010). Haunting experiences: An integrative approach. In M. Smith (Ed), *Anomalous experiences: Essays from parapsychological and psychological perspectives* (pp. 108-119). Jefferson, NC: McFarland & Co.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776-783. doi.org/10.1037/h0043424

Palmer, J. (1974). A case of RSPK involving a ten-year-old boy: The Powhatan poltergeist. *Journal of the American Society for Psychical Research*, *68*, 1-33.

Parra, A. (2018). Perceptual-personality variables associated with entity encounter experiences. *Australian Journal of Parapsychology*, *18*, 23-48.

Parra, A. & Argibay, J.C. (2016). The boundary construct and anomalous experiences in psychics. *Journal of the Society for Psychical Research, 80*, 13-23.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825-2830.

Persinger, M. A., & Cameron, R. A. (1986). Are earth faults at fault in some poltergeist-like episodes? *Journal of the American Society for Psychical Research*, *80*, 49-73.

Playfair, G. L. (1980). *This house is haunted: An investigation of the Enfield poltergeist*. London, England: Souvenir Press.

Potts, J. (2004). Ghost hunting in the twenty-first century. In J. Houran (Ed.), *From shaman to scientist: Essays on humanity's search for spirits* (pp. 211-232). Lanham, MD: Scarecrow Press.

Pratt, J. G., & Palmer, J. (1976). An investigation of an unpublicized family poltergeist [Abstract]. In J. D. Morris, W. G. Roll, & R. D. Morris (Eds.), *Research in parapsychology 1975* (pp. 109-115). Metuchen, NJ: Scarecrow Press.

Preti, A., Vellante, M., & Petretto, D. R. (2017) The psychometric properties of the "Reading the Mind in the Eyes" test: An item response theory (IRT) analysis. *Cognitive Neuropsychiatry*, *22*, 233-253. doi.org/10.1080/13546805.2017.1300091

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. (Copenhagen, Danish Institute for Educational Research), expanded edition (1980) with foreword and afterword by B. D. Wright. Chicago, IL: University of Chicago Press.

Roll, W. G. (1977). Poltergeists. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 382-413). New York: Van Nostrand Reinhold.

Rugg, H. (1963). *Imagination*. New York, Evanston, and London, England: Harper & Row.

Schmeidler, G. R. (2001). Afterword. In J. Houran & R. Lange (Eds.), *Hauntings and poltergeists: Multidisciplinary perspectives* (pp. 307-309). Jefferson, NC: McFarland & Co.

Schofield, M. B., Baker, I. S., Staples, P., & Sheffield, D. (2018). Creation and validation of the Belief in the Supernatural Scale. *Journal of Parapsychology*, *82*, 41-64. doi.org/10.30891/jopar.2018.01.04

Schofield, K., & Claridge, G. (2007). Paranormal experiences and mental health: Schizotypy as an underlying factor. *Personality and Individual Differences*, *43*, 1908-1916. doi.org/10.1016/j.paid.2007.06.014

Sen, M., & Yesilyurt, E. (2014). The development of Paranormal Belief Scale (PBS) for science education in the context of Turkey. *International Journal of Education in Mathematics, Science and Technology*, *2*, 107-115. doi.org/10.18404/ijemst.40479

Sinclair, M., & Ashkanasy, N. M. (2005). Intuition: myth or decision-making tool? *Management Learning*, 36, 353-370. doi.org/10.1177/1350507605055351

Slosson, E. E. (1899). Shorter communications and discussions: A lecture experiment in hallucinations. *Psychological Review*, *6*, 407-408. doi.org/10.1037/h0071184

Stokes, D. M. (2017a). White crows rising: Using spontaneous cases to establish psi. *Australian Journal of Parapsychology*, *17*, 47-60.

Stokes, D. M. (2017b). A murder of white crows: Additional cases of unequivocal spontaneous psi. *Australian Journal of Parapsychology*, *17*, 171-186.

Storm, L., Drinkwater, K., & Jinks, A. L. (2017). A question of belief: An analysis of item content in paranormal belief questionnaires. *Journal of Scientific Exploration, 31*, 187-230.

Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, *55*, 293-326. doi.org/10.1007/BF02294821

Tobacyk, J. J. (2004). A Revised Paranormal Belief Scale. *International Journal of Transpersonal Studies*, *23*, 94-98. doi.org/10.24972/ijts.2004.23.1.94

van der Linden, W.J. & Hambleton, W.J. (1997). *Handbook of modern item response theory*. New York: Springer. doi.org/ 10.1007/978-1-4757-2691-6

Werner, H. (1948). *Comparative psychology of mental development*. New York: International Universities Press.

Werner, H. (1934/1978). Unity of the senses. In S. S. Barten & M. B. Franklin (Eds.), *Developmental processes: Heinz Werner's selected writings, vol. 1* (pp. 153-167). New International Universities Press.

Wessely, S. (1987). Mass hysteria: Two syndromes? *Psychological Medicine*, *17*, 109-120. doi.org/10.1017/S0033291700013027

Wessely, S. (2000). Responding to mass psychogenic illness. *New England Journal of Medicine*, *342*, 129-130. doi.org/10.1056/NEJM200001133420212

Williams, B., & Ventola, A. (2011). Poltergeist phenomena: A primer on parapsychological research and perspectives. Public Parapsychology.org. Retrieved from: publicparapsychology.org/Public%20Parapsych/Poltergeist%20Phenomena%20Primer%20Final.pdf

Wiseman, R., Watt, C., Greening, E., Stevens, P., & O'Keeffe, C. (2002). An investigation into the alleged haunting of Hampton Court Palace: Psychological variables and magnetic fields. *Journal of Parapsychology, 66*, 387-408.

Wright, B. D., & Douglas, G. (1975). *Best test design and self-tailored testing*. MESA Memorandum No. 19. Department of Education, University of Chicago, Illinois.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.

Wright, B. D., & Mok, M. (2000). Understanding Rasch measurement: Rasch models overview. *Journal of Applied Measurement*, *1*, 83-106.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: MESA Press.

Zausner, T. (1998). *Creativity, panic attacks, and states of awe: A model from nonlinear dynamics.* Paper presented at the Eight Annual Conference of the Society of Chaos Theory in Psychology and the Life Sciences, Boston, MA. July 28 - August 1. doi.org/10.1037/e631962011-001

Appendix A

Rasch Scoring Table for the Survey of Strange Events (SSE)

Raw Sum	Scale Score	Standard Error of Measurement
0[a]	22.3	11.7
1	30.3	6.6
2	35.2	4.8
3	38.4	4.1
4	40.7	3.7
5	42.7	3.4
6	44.4	3.2
7	45.9	3.0
8	47.3	2.9
9	48.6	2.8
10	49.8	2.7
11	51.0	2.7
12	52.1	2.6
13	53.2	2.6
14	54.3	2.6
15	55.3	2.6
16	56.4	2.6
17	57.5	2.6
18	58.5	2.6
19	59.6	2.6
20	60.7	2.7
21	61.9	2.7
22	63.0	2.7
23	64.3	2.8
24	65.6	2.9
25	67.0	3.0
26	68.5	3.2
27	70.2	3.4
28	72.2	3.7
29	74.6	4.1
30	77.8	4.9
31	82.8	6.6
32[a]	90.9	11.7

[a] Extreme values are approximations only.

Appendix B

The Survey of Strange Events

Instructions

- This survey asks about unusual experiences you may have had at a specific place (indoors or outdoors) that people might claim is haunted by a ghost or other paranormal force. Or, your experiences might have happened at a location with no paranormal reputation. Please answer this survey with only one particular location in mind.
- Carefully read the list of "strange" events below and think about the location you visited. Next, write TRUE (or 'T') next to any of the events that you have experienced at this location. When a question says "mysterious" it means that the event had no obvious or immediate explanation from your point of view.
- Your important information is confidential, so please be honest and detailed.

• **Your Age:_____** • **Your Gender:_____** • **Your Country of Origin:_____** • **How long ago did your unusual experiences at this location occur?** • **(round to nearest month):_____** • **In what type of location did the experiences occur (e.g., house, park, cemetery, pub, office, etc.)? _____**	**TRUE (T) or FALSE (F)**
Personal Experiences	
1. I saw with my naked eye a non-descript visual image, like fog, shadow or unusual light	
2. I saw with my naked eye an "obvious" ghost or apparition – a misty or translucent image with a human form	
3. I saw with my naked eye an "un-obvious" ghost or apparition – a human form that looked like a living person	
4. I smelled a mysterious odor that was *pleasant*	
5. I smelled a mysterious odor that was *unpleasant*	
6. I had a *positive* feeling for no obvious reason, like happiness, love, joy, or peace	
7. I had a *negative* feeling for no obvious reason, like anger, sadness, panic, or danger	
8. I felt odd sensations in my body, such as dizziness, tingling, electrical shock, or nausea (sick in my stomach)	
9. I had a mysterious taste in my mouth	
10. I felt guided, controlled or possessed by an outside force	

11. I saw beings of divine or evil origin, such as angels or demons	
12. I saw folklore-type beings that were not human, such as elves, fairies, or other types of "little people"	
13. I communicated with the dead or other outside force	
14. I had the mysterious feeling of being watched, or in the presence of an invisible being or force	
15. I had a sense of déjà vu, like something was strangely familiar to me about my thoughts, feelings or surroundings	
Physical Events	
16. I heard mysterious sounds that could be recognized or identified, such as ghostly voices or music (with or without singing)	
17. I heard mysterious "mechanical" or non-descript noises, such as tapping, knocking, rattling, banging, crashing, footsteps or the sound of opening/closing doors or drawers	
18. I heard on an audio recorder mysterious sounds that could be recognized or identified, such as ghostly voices or music (with or without singing)	
19. I heard on an audio recorder mysterious "mechanical" or non-descript noises, such as tapping, knocking, rattling, banging, crashing, footsteps or the sound of opening/closing doors or drawers	
20. I felt a mysterious area of *cold*	
21. I felt a mysterious area of *heat*	
22. I experienced objects disappear or reappear around me	
23. I saw objects moving on their own across a surface or falling	
24. I saw objects flying or floating in midair	
25. Electrical or mechanical appliances or equipment functioned improperly or not at all, including flickering lights, power surges or batteries "going dead" in electronic devices (e.g., camera, phone, etc.)	
26. Pictures from my camera or mobile device captured unusual images, shapes, distortions or effects	
27. Plumbing equipment or systems (faucets, disposal, toilet) functioned improperly or not at all	
28. I saw objects breaking (or discovered them broken), like shattered or cracked glass, mirrors or housewares	
29. I felt a breeze or a rush of wind or air, like something invisible was moving near me	
30. Fires have started mysteriously	

31. I was mysteriously touched in a *non-threatening* manner, like a tap, touch or light pressure on my body	
32. I was mysteriously touched in a *threatening* manner, such as a cut, bite, scratch, shove, burn or strong pressure on my body	

Quantifier la Phénoménologie des Épisodes Fantomatiques : 2e Partie – Un Modèle Rasch des Témoignages Spontanés

Résumé. En utilisant un échantillon de témoignages « spontanés » auto-rapportés (apparemment sincères et sans biais d'amorçage, *N = 426*), nous avons calibré le « questionnaire des événements étranges » (SSE) en 32-items selon la méthode de Rasch, afin de quantifier la phénoménologie des épisodes fantomatiques tout en évaluant les biais de réponse relatifs à l'âge ou au genre des répondants. L'inventaire inclut des expériences psychologiques typiques de la hantise, et des manifestations physiques communes aux perturbations de type poltergeist. Les résultats supportent les précédentes suggestions selon lesquelles les témoignages « spontanés » montrent un pattern comportemental (cumulatif) prédictible, ainsi qu'une structure factorielle unidimensionnelle. De plus, comparativement aux récits spontanés, nous avons identifié de forts biais de réponses sur le SSE sur quatre conditions de contrôle (à savoir le style de vie, l'amorce, la fantaisie, et l'illicite). La modélisation statistique prédit avec succès les appartenances de groupe, corroborant le fait que les expériences spontanées diffèrent systématiquement des « imposteurs » de certaines manières. Le SSE est un outil de mesure robuste de l'intensité Générale des épisodes fantomatiques (fiabilité Rasch = 0.87) et peut servir pour opérationnaliser, de façon standardisée, les anomalies spécifiques dans les sondages, les études de terrain, et les investigations qui codent des données en réponse libre ou du matériel issu de cas spontanés à des fins d'analyse quantitative.

Quantifizierung der Phänomenologie geisterhafter Episoden: Teil II - Ein Rasch-Modell über Spontanberichte

Zusammenfassung. Anhand einer Stichprobe von selbsterlebten "spontanen" Berichten (mutmaßlich aufrichtig und unbeeinflusst, N = 426) eichten wir einen aus 32 Items bestehenden Rasch-basierten "Survey of Strange Events (SSE)", um die Phänomenologie geisterhafter Episoden zu quantifizieren und gleichzeitig Antworttendenzen in Bezug auf Alter oder Geschlecht der Berichterstatter einzuschätzen. Die Umfrage umfasste psychologische Erfahrungen, die typisch für Geistererscheinungen sind, und physikalische Vorfälle, die bei spukähnlichen Vorfällen auftreten. Die Ergebnisse unterstützten frühere Vorstellungen, dass "spontane" Berichte ein vorhersagbares (kumulatives) Verhaltensmuster haben und eine eindimensionale Faktorenstruktur aufweisen. Darüber hinaus identifizierten wir im Vergleich zu Spontanberichten starke Antwortverzerrungen der SSE über vier Kontrollbedingungen hinweg (Lebensstil, Erwartet, Phantasie und Unerlaubt). Eine statistische Modellierung sagte erfolgreich Gruppenzugehörigkeiten mit guter Genauigkeit voraus, was bestätigt, dass sich spontane Erfahrungen in bestimmter Weise systematisch von "Betrügern" unterscheiden. Die SSE ist ein robustes Maß für die Gesamtintensität geisterhafter Episoden (Rasch-Reliabilität = 0,87) und dient zur standardmäßigen Operationalisi-

erung spezifischer Anomalien bei Umfragen, Feldstudien und Untersuchungen, die Free-Response-Daten oder spontanes Fallmaterial für die quantitative Analyse kodieren.

Cuantificación de la Fenomenología de los Episodios Fantasmales: Parte II - Un Modelo Rasch de Relatos Espontáneos

Resumen. Utilizando una muestra de relatos “espontáneos” autoinformados (aparentemente sinceros y no influenciados, *N* = 426), calibramos una “Survey of Strange Eventos (SSE; o Encuesta de Eventos Extraños)” basada en Rasch de 32 ítems para cuantificar la fenomenología de los episodios fantasmales al mismo tiempo que evaluamos sesgos de respuesta relacionados con la edad o el género de los participantes. Este inventario incluyó experiencias psicológicas típicas de experiencias de casas encantadas y manifestaciones físicas comunes a perturbaciones semejantes al poltergeist. Los resultados respaldaron sugerencias anteriores de que los relatos “espontáneas” tienen un patrón de comportamiento (acumulativo) predecible y muestran una estructura factorial unidimensional. Además, en comparación con los relatos espontáneos, identificamos fuertes sesgos de respuesta en la ESS en cuatro condiciones de control (Estilo de vida, Influído, Fantasía, e Ilícito). El modelado estadístico predijo la pertenencia a grupos con buena precisión, corroborando que las experiencias espontáneas difieren sistemáticamente en ciertos aspectos de los “impostores.” El SSE es una medida robusta de la intensidad general de los episodios fantasmales (fiabilidad Rasch = 0.87) y puede servir como una operacionalización estándar de anomalías en encuestas, estudios de trabajo de campo, e investigaciones que codifiquen datos de respuesta libre o material de casos espontáneos para análisis cuantitativos.