Five degrees of (non)sense: Investigating the connection between bullshit receptivity and susceptibility to semantic illusions

Dario Paape*

Abstract. Individual differences in people's tendency to see bullshit statements such as *Perceptual reality transcends subtle truth* as meaningful and possibly profound have become an active topic of research in judgment and decision making in recent years. However, (psycho)linguistics has so far paid little attention to the topic, despite its obvious appeal for language processing research. I present an experiment that investigated possible shared traits contributing to individual bullshit receptivity and susceptibility to semantic illusions, which occur when compositionally incongruous sentences receive plausible but unlicensed interpretations (e.g., *More people have been to Russia than I have*). The results show relatively little indication of an individual-level tendency to both fall for bullshit and for linguistic illusions. Implications for future psycholinguistic research into bullshit processing are discussed.

Keywords. bullshit, semantic illusion, experimental semantics, individual differences

1. Introduction. Consider the three sentences in (1), which exemplify different types of *bullshit* or *obscurantism* (Buekens & Boudry 2015).[1]

(1) a. Hidden meaning transforms unparalleled abstract beauty.
b. A cyclic ionization whose only final result is to induce plasmatic solubility is impossible.
c. Deliberate ambivalence is inherent to the approach, yielding qualities where things convulse and stutter in emerging vitality.

(1-a) is an example "pseudo-profound" or "pseudo-transcendental" bullshit (Pennycook et al. 2015, Čavojová et al. 2020), (1-b) is an example of "scientific" bullshit (Evans et al. 2020), and (1-c) is an example of "International Art English" (Turpin et al. 2019, Rule & Levine 2012). The common feature of these three utterances is that they are *unclarifiably unclear* (Cohen 2002): It's impossible to restate in plain language what the sentences mean, if they mean anything at all. They contain a multitude of abstract terms, jargon, and buzzwords, which are "put together randomly in a sentence that retains syntactic structure" (Pennycook et al. 2015; p. 548).

Bullshit is often produced when speakers talk about things they have no detailed knowledge of, or about things that lack an established "ground truth" (such as art or philosophy; Frankfurt 1986). Yet, despite its apparent vacuousness, bullshit is sometimes perceived as being meaningful or even profound.[2] Readers vary in their *bullshit receptivity*, that is, in their tendency to imbue bullshit with meaning. Iacobucci & Cicco (in press) present a review of 40 studies published between 2015 and 2021 that are concerned with individual differences in bullshit receptivity. Among the

*Author: Dario Paape, University of Potsdam (paape@uni-potsdam.de).

[1]Buekens & Boudry (2015) draw a contrast between bullshit and obscurantism that is based on the utterer's intention, but under the content-based characterization of bullshit that I will adopt, the contrast becomes irrelevant.

[2]Meaningfulness is, of course, subjective to some extent, and apparent nonsense may be argued to sometimes contain deep wisdom (Dalton 2016). However, this is especially unlikely for sentences that are intentionally generated as bullshit (Pennycook et al. 2016).

factors investigated are intelligence (Bainbridge et al. 2019), individual proclivity towards analytic versus intuitive thinking (e.g. Pennycook et al. 2016), and self-regulation ability (Petrocelli et al. 2020).

For the present investigation, I will focus on two additional traits: *Interpretive charity*, that is, a person's tendency to assume that statements are meaningful and true by default (e.g., Sperber 2010), and illusory pattern perception or *apophenia*, that is, a person's tendency to see patterns where none exist (DeYoung et al. 2012). An indiscriminate bias to judge sentences as meaningful and profound has been argued to contribute to pseudo-profound bullshit receptivity in particular (Pennycook et al. 2015), though the ability to *discriminate* between actual profundity and bullshit appears to be a stronger driver of individual differences (Bainbridge et al. 2019).

Apophenia or illusory pattern perception as a driver of bullshit receptivity has been investigated by Walker et al. (2019) and by Bainbridge et al. (2019). Walker et al. (2019) presented participants with random sequences of 10 coin tosses (e.g., HTHHTTTTHH) and asked them to indicate on a 1–7 scale how strongly they felt that the results were random or pre-determined. In another experiment, participants were shown images of objects embedded in visual noise, as well as images containing only visual noise, and were asked whether the image contained an object. Profundity ratings for pseudo-profound bullshit were positively correlated both with the coin-toss determinism measure and with the tendency to perceive objects in pure noise. Bainbridge et al. (2019) used a more indirect measure of apophenia, by asking participants about various paranormal and otherwise unusual beliefs ("It is possible to move material objects with only one's thoughts"). Paranormal beliefs correlate with illusory pattern perception (van Prooijen et al. 2018). Bainbridge et al. (2019) found a positive correlation between apophenia and pseudo-profound bullshit receptivity, in line with Walker et al.'s findings.

To my knowledge, despite interesting implications for semantic and pragmatic processing, bullshit has not received much attention in psycholinguistics.[3] The fact that some people find the sentences in (1) meaningful suggests a role for processes that create meaning independently of the compositional makeup of the utterance and the lexical meanings of the words, to the extent that the average reader even has access to the lexical meaning of a word like *ionization*.

However, there is a related class of phenomena that *has* been studied in psycholinguistics: semantic illusions. A semantic illusion occurs when a semantically incongruous and arguably meaningless sentence appears well-formed and sensible. The sentences in (2) are examples of such illusions.

(2) a. More people have been to Russia than I have.
b. No head injury is too trivial to be ignored.

Sentence (2-a) is an instance of the so-called *comparative illusion* (CI). The sentence is compositionally incongruous because an amount of people (*more people than X*) is compared with an event (*I have [been to Russia]*), which should not yield a sensible meaning. Yet, the anomaly of CI sentences is often not consciously detected, and the sentences are rated as relatively acceptable (e.g., O'Connor 2015, Wellwood et al. 2018). The illusion is not random, however: CI sentences are rated more highly when they contain repeatable rather than unrepeatable events (?*More girls*

[3]The only published source I could find in which a link is made is de Almeida (2018).

graduated high school than the boy did), suggesting a preference for an illicit event comparison reading (Wellwood et al. 2018, De Dios-Flores 2016), but also when they contain semantically plural object NPs (*More cats have mouse toys than the dog does*; O'Connor 2015), suggesting the availability of an equally illicit cardinality comparison (*more mouse toys*). In both cases, the ungrammatical sentence is apparently coerced into an interpretable form.

Sentence (2-b) is a so-called *depth charge* (DC) sentence (e.g., Sanford & Sturt 2002). Like CI sentences, DC sentences are compositionally incongruous: The degree phrase *too trivial to be ignored* presupposes that more trivial things are *less* likely to be ignored, which runs contrary to world knowledge (compare *too trivial to be treated*). Furthermore, (2-b) is usually interpreted to mean *Don't ignore head injuries*, but the compositional meaning is *Ignore all head injuries* (compare *No landmine is too small to be banned*; Wason & Reich 1979). Different theories have attributed the DC illusion to superficial processing (Wason & Reich 1979, Paape et al. 2020), to unconscious repair of an assumed speech error (Zhang et al. 2022), or to the availability of a stored grammatical construction with an idiosyncratic meaning (Fortuin 2014, Cook & Stevenson 2010). As for CI sentences, mechanisms beyond standard compositional semantics must be recruited in order to make sense of the construction.

Both CI sentences and DC sentences show considerable variability in judgments between speakers (Wellwood et al. 2018, Leivada 2020, Paape et al. 2020). Some speakers tend to almost always find illusion sentences acceptable while others categorically reject them. However, to my knowledge, a possible shared susceptibility to different semantic illusions within the same speaker has never been investigated. Nevertheless, it is plausible that such general differences in susceptibility exist. Both for CI and for DC sentences, it has been suggested that there is a threshold of processing complexity that differs between speakers, and that processing is aborted once this threshold is reached and the meaning representation is deemed "good enough" (Paape et al. 2020, Paape 2021, Leivada 2020). Low threshold settings can result in failures to notice linguistic anomalies and cause acceptability illusions for malformed sentences (e.g., Christianson 2016). This perspective is in line with the proposal that readers have individual *standards of coherence* that affect their depth of processing during reading (van den Broek et al. 2011, 2001), with lower standards of coherence meaning less processing.

Which brings us back to bullshit: It would appear that in order to accept a bullshit statement such as *Hidden meaning transforms unparalleled abstract beauty* as meaningful, one would need to set a relatively low standard of coherence, because the more one thinks about the semantic content of the sentence, the less meaningful it becomes. Anecdotally, the same is often true for CI and DC illusion sentences. At the same time, however, both bullshit sentences and illusion sentences require *enrichment* of the linguistic input to be perceived as meaningful:[4] Readers presumably *want* the sentences to make sense, and thus project plausible meanings into them. Superficial processing on the one hand does not necessarily contradict semantic enrichment on the other: Readers may simply shift their focus away from the actual structure of the input and onto other, stimulus-independent sources of meaning, such as their pre-existing knowledge or opinions (Paape 2021). Highly apophenic individuals with a "greater tendency to go beyond the available data" (Walker et al. 2019; p. 111) and to "creat[e] meaning where no meaning exists" (ibid., p. 117)

[4]See de Almeida (2018) for a similar point.

may be especially prone to this form of enrichment. The prediction is thus that the perceived meaningfulness of bullshit sentences, CI sentences and DC sentences should covary jointly within speakers, and that highly apophenic readers should be especially likely to see meaning in all three sentence types.

But how can individual differences in apophenia be distinguished from differences in interpretive charity – or are they ultimately the same thing? There are reasons to assume that they are not. Illusory pattern perception, which is at the core of aponenia, is triggered by cues that at least *resemble* a pattern (Walker et al. 2019). Thus, the more a given sentence *resembles* a well-formed utterance, the more apophenic interpretation should occur. Crucially, bullshit sentences and illusion sentences are not *nonsense* in the sense that they are random jumbles of words.[5] On some level, these sentences *look* and *feel* "good enough" to pass inspection. Interpretive charity, on the other hand, could plausibly make *anything* appear meaningful: a highly charitable reader may force even "word salad" to have meaning (Fowler 1969), just like any artwork can be imbued with meaning by a charitable viewer.

In what follows, I present a web-based experiment that investigated possible correlations between bullshit receptivity, susceptibility to semantic illusions, and apophenia. Interpretive charity was controlled for by also including sensible sentences and nonsense sentences in the experiment. The experiment used German sentences, and was run with a sample of German native speakers.

2. Experimental study. The experiment deviated from previous studies on bullshit by having participants judge the meaningfulness of the sentences instead of their profundity (e.g., Pennycook et al. 2015) or truthfulness (e.g., Evans et al. 2020). This more basic level of judgment was chosen to allow comparison with the illusion sentences, which are not intended to be profound or necessarily true. Three types of bullshit were included: pseudo-profound bullshit, scientific bullshit, and International Art English. Pseudo-profound bullshit receptivity has been found to correlate both with scientific bullshit receptivity (Evans et al. 2020) and with receptivity towards International Art English (Turpin et al. 2019), suggesting shared underlying traits.

In order to somewhat offset participants' interpretive charity and to naturalize the judgment of unusual sentences, participants were told to imagine that the sentences had been generated by an AI system, and that they were helping to improve the system by distinguishing between "good" and "bad" sentences. The pattern recognition task intended to measure apophenia was also embedded in the AI scenario (see below).

2.1. PARTICIPANTS. 100 self-identified German native speakers were recruited on Prolific (https://www.prolific.co; Palan & Schitter 2018). They were paid £3,50 for their participation.[6]

[5]For an illuminating discussion of different kinds of nonsense, see Diamond (1981).

[6]Because Prolific is based in the UK, renumeration is calculated in British pounds.

2.2. MATERIALS. Example items for each experimental condition are shown in (3).[7]

(3) a. **Bullshit**

The invisible is beyond new timelessness. (pseudo-profound)

Energy can deteriorate based on closed-circuit alliterations of an afocal system. (scientific)

The banality of some gestures is disconcerting, and in their strangeness, they convey a future created in the past. (International Art English)

b. **Comparative illusion (CI)**

More spectators in the theater are proud Americans than the actor is.

c. **Depth charge illusion (DC)**

In the end, you realize that no plan is too unrealistic to be scrapped.

d. **Sensible**

Your teacher can open the door, but you must enter by yourself. (profound)

Newborns require constant attention. (mundane)

e. **Nonsense**

One can say that flowers with a lot of old nettles do not limp without great experience value.

Each participant rated 24 bullshit sentences in total (8 of each type), as well as 16 CI sentences and 12 DC sentences.[8] These were randomly intermixed with 24 sensible sentences (12 profound, 12 mundane) and 20 nonsense sentences. The nonsense sentences were generated in a stream-of-consciousness manner by the author and validated as being nonsense by three German speakers.

The 20 stimuli for the pattern recognition task were scatterplots created in R (R Core Team 2022). For each plot, 20 random floating-point numbers between 0 and 30 were generated for both the X and the Y coordinate using R's runif() function. Examples are shown in Figure 1.

Participants with high apophenia were expected to "detect" more patterns in the scatterplots, and possibly show longer response times, because they may spend more time trying to find patterns.

2.3. PROCEDURE. Participants provided informed consent prior to experimentation. The AI scenario was first introduced. Participants were told that they should rate the meaningfulness of each sentence in comparison to a completely meaningless baseline sentence (*Bats don't go defiantly into the computer next to love without pumping*). This baseline sentence was presented in each trial. The rating scale for the target sentences ranged from 1 ("equally bad") to 7 ("much better"). Participants could freely reread the sentences as many times as they wished, but were instructed not to "overanalyze" the sentences and to rely more on their linguistic intuition. Reading times in

[7]Pseudo-profound bullshit sentences, scientific bullshit sentences and International Art English sentences were translated and adapted from Pennycook et al. (2015), Evans et al. (2020), and Turpin et al. (2019). Comparative illusion sentences were translated and adapted from O'Connor (2015). Depth charge sentences were adapted from Paape et al. (2020). Sensible sentences were mostly translated and adapted from Pennycook et al. (2015), but also featured some new additions.

[8]A total of 24 DC sentences were created, but only 12 were presented to each participant to limit the duration of the experimental session and to prevent carryover effects.

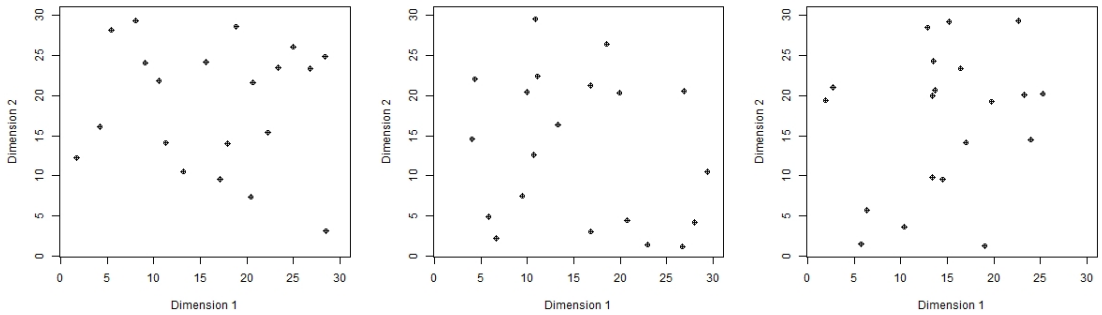

Figure 1: Stimuli used for the pattern recognition task.

milliseconds were recorded in addition to the ratings.

After all sentences had been rated, the pattern recognition task was administered. Participants were presented with the scatterplots in random order and were told that these represented the "neuronal activations" of the AI model, which are sometimes random and sometimes structured. They were instructed to give a binary judgment of whether they intuitively felt that each pattern was structured as opposed to random. Participants were also told that all patterns may be random, or all patterns may be structured.

2.4. DATA ANALYSIS. The statistical analysis was carried out using the Stan language for Bayesian inference (Stan Development Team 2022). Because the rating data are ordinal, they were analyzed with a cumulative logit model. This kind of model assumes a continuous latent variable underlying the Likert scale ratings, together with a set of thresholds or "cutpoints" that group the latent values into the rating "bins" enforced by the discrete scale (Liddell & Kruschke 2018). Because the individual differences on the *latent* scale are of interest, and because participants may differ in their use of the discrete scale, recovering the underlying continuous values by modeling participants' individual cutpoints is of major importance (see below for the implementation).

Reading times were analyzed assuming a lognormal likelihood. For both dependent variables, a hierarchical model with correlated "random" (or varying) intercepts and slopes by participants and by items was fitted (e.g., Pinheiro & Bates 2000, Gelman & Hill 2007). This means that an individual adjustment to the population-level estimate was estimated for each subject and for each item in order to account for the non-independence of measurements and to capture inter-individual differences (Barr et al. 2013). In the present study, the *correlations* between these adjustments, which are also estimated from the data, are of major interest: The question is whether someone who gives higher-than-average ratings to bullshit sentences will, for instance, also give higher-than-average ratings to comparative illusion sentences.

In Stan, it is possible to set up a model that estimates a separate average (intercept) value for each sentence type, along with participant- and item-specific adjustments, and, crucially, correlations between the adjustments *across* sentence types. A shortened and simplified notation of the model is shown in equation 1. The model code and data are available at `https://osf.io/54wh7`.

For each trial n, the rating comes from an ordered logistic distribution with a vector C of six cutpoints. The values $\gamma_{1...6}$ demarcate the population-level boundaries between rating "bins" in logit space, which are adjusted for each participant in order to account for differences in the use of the Likert scale. A population-level intercept α on the latent scale is estimated for each sentence type and adjusted by participant (u_i) and by item (w_j). The adjustments follow a normal distribution with a to-be-estimated standard deviation σ_u, σ_w. The correlations between the adjustments $u_{1...m}$ and w can be recovered from the estimates of the variance-covariance matrix of the random effects. Regularizing priors were used for all parameters, including LKJ priors (Lewandowski et al. 2009) with η set to 2 for the correlations.

$$
\begin{aligned}
\mu_{1,n} &= \alpha_1 + u_{1,i} + w_j && (bullshit) \\
\mu_{2,n} &= \alpha_2 + u_{2,i} + w_j && (comparative) \\
\mu_{3,n} &= \alpha_3 + u_{3,i} + w_j && (depth\ charge) \\
C_{1,n} &= \gamma_1 + u_{4,i} \\
C_{2,n} &= \gamma_2 + u_{5,i} \\
&\dots \\
u_{1,i} &\sim Normal(0, \sigma\_u_1) \\
&\dots \\
w_j &\sim Normal(0, \sigma\_w) \\
Rating_n &\sim Ordered\_logistic(\mu_{x,n}, C_n)
\end{aligned}
\tag{1}
$$

Participants' reading times as well as their reaction times and judgments for the scatterplots were also analyzed within the same model.[9] This approach allows for the direct estimation of the critical random-effects correlations across measures, without any pre-aggregration of the data, and thus without any loss of information. Furthermore, the addition of item-specific adjustments allows for better estimation of the subject-level individual differences: If a given participant sees a pattern in a scatterplot that has an above-average "pattern-likeness" across all participants, this is much less informative than if the participant sees a pattern in a scatterplot with a below-average "pattern-likeness".

For the reading times, a slope for the number of characters in the sentence was added to the model, along with adjustments by subject. The slope adjustment for each participant was intended as a measure of their tendency towards superficial or "good enough" processing. It has been found that word length effects are reduced during "mindless" reading (Schad et al. 2012, Reichle et al. 2010, Franklin et al. 2011), so it is plausible that the less attention a participant is paying to the sentences, the less the overall sentence length should affect their reading times.

2.5. RESULTS. Table 1 shows mean ratings and standard deviations of the by-subject means by sentence type.

[9]It is advantageous for this approach to have the dependent variables on similar scales. Here, the binary judgments and the sentence ratings are analyzed on the log-odds (logit) scale while reading times and reaction times are analyzed on the log scale.

	$\overline{Rating}$	SD
bullshit	4.37	0.85
comparative	4.42	1.11
depth charge	5.08	0.92
sensible	6.58	0.52
nonsense	2.05	0.77

Table 1: Mean ratings and standard deviations of by-subject means by sentence type.

Unsurprisingly, sensible sentences were rated the most meaningful in comparison to the baseline sentence while nonsense sentences were rated the least meaningful. The three other sentence types fall in between, with somewhat higher ratings for DC sentences than for bullshit and CI sentences. CI sentences showed the greatest rating variability between subjects, followed by depth charge sentences. Sensible sentences showed the lowest variability.

Figure 2 shows the correlation estimates for the participant-level random effects extracted from the Stan model, along with the associated 95% highest density intervals (HDIs) of the posterior distributions (e.g., Kruschke 2014) computed using the bayestestR package (Makowski et al. 2019).

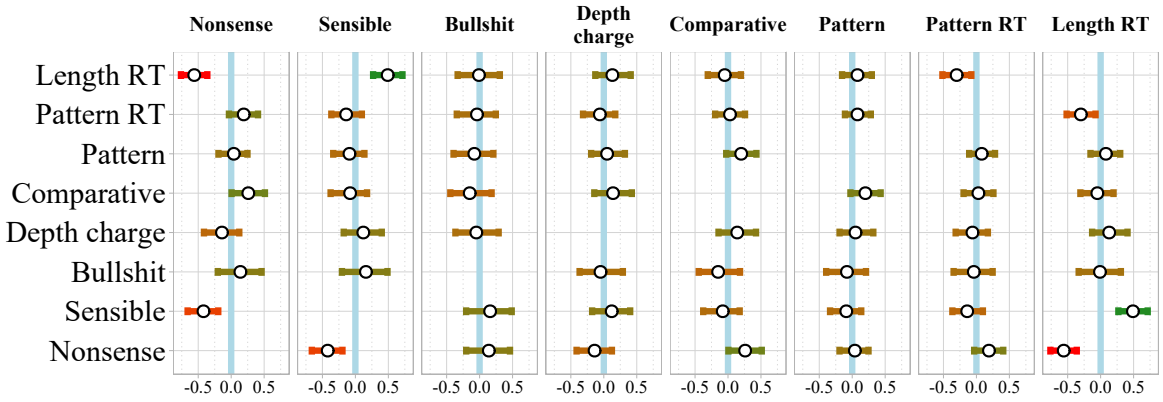

Figure 2: 95% HDIs of the correlation estimates for the participant-level random effects. The diagonal is not plotted because the correlation of each adjustment with itself is trivially 1. Colors show the sign and magnitude of the correlation estimate (red = negative, green = positive).

Focusing first on the main prediction, namely a positive correlation between apophenia, as measured by increased pattern spotting and longer reaction times for the scatterplots, and the perceived meaningfulness of bullshit sentences and illusion sentences, the data do not show any strong indication of the predicted relationship. Only the perceived meaningfulness of comparative illusion sentences shows some indication of a positive correlation with pattern spotting at the participant level (95% HDI: $[-0.03, 0.43]$). Furthermore, the HDIs of the pairwise correlations between by-participant adjustments for bullshit meaningfulness, CI meaningfulness and DC meaningfulness are centered around values close to zero, indicating no evidence for the predicted positive correlation due to individual differences in interpretive charity.

There is a negative correlation at the participant level between the perceived meaningfulness of sensible sentences and that of nonsense sentences (HDI: [−0.67, −0.19]). Perceived meaningfulness of nonsense is negatively correlated with the effect of sentence length on reading times (HDI: [−0.76, −0.36]), while perceived meaningfulness of sensible sentences is positively correlated with the sentence length effect (HDI: [0.27, 0.7]). This suggests that more superficial readers, who were less affected by sentence length, perceived nonsense as more meaningful and sensible sentences as less meaningful than average. Superficial or inattentive readers thus seem less able or willing to distinguish between the sentence types, compared to more attentive readers.

Finally, participants who spent more time looking at the scatterplots showed smaller effects of sentence length on reading time (HDI: [−0.52, −0.08]), as well as higher perceived meaningfulnesss of nonsense sentences (HDI: [−0.03, 0.41]). These correlations can be seen as tentative evidence for a connection between apophenia and superficial language processing. However, there is no indication in the data that individuals who spent more time *looking* for patterns in the scatterplots ended up "finding" more of them.

3. Discussion. The goal of this study was to investigate individual differences in the processing of bullshit statements (*Hidden meaning transforms unparalleled abstract beauty*) and two types of semantic illusion, namely the comparative illusion (CI; *More people have been to Russia than I have*) and the depth charge illusion (DC; *No head injury is too trivial to be ignored*). The underlying intuition was that readers who find meaning in bullshit statements may also find meaning in semantic illusion sentences. Based on the existing bullshit literature, two traits were hypothesized to contribute to both bullshit receptivity and illusion receptivity: The first is a person's general bias towards finding statements meaningful or even profound, irrespective of content, that is, their tendency towards interpretive charity (e.g. Pennycook et al. 2015). The second is a person's tendency to actively *create* meaning by seeking patterns for which there is no objective evidence in the data, that is, their tendency towards apophenia (Walker et al. 2019).

Statistically, shared individual differences in interpretive charity, apophenia, and the perceived meaningfulness of bullshit and semantic illusions were explored by looking at the correlations of subject-level random effects in a hierarchical model. However, there was little evidence in the data that would support the hypothesized connections. There was no indication that people who found bullshit sentences more meaningful also found illusion sentences more meaningful, or that finding CI sentences meaningful correlates with finding DC sentences meaningful. Regarding effects of apophenia, only the perceived meaningfulness of CI sentences but not that of bullshit sentences or DC sentences showed a positive correlation with finding patterns in random scatterplots, thus casting doubt on a shared underlying mechanism of "meaning creation".

The absence of evidence for a connection between apophenia and bullshit endorsement is unexpected given earlier findings by Walker et al. (2019). However, the mismatch may be due to a variety of factors: the smaller participant sample, the use of a mixture of bullshit "genres" in the current study, and/or differences between the patttern-spotting tasks. For instance, unlike the one used by Walker et al. (2019), the pattern spotting task used in the current study did not have a baseline in which *real* patterns needed to be identified. Furthermore, the current experiment used judgments of meaningfulness whereas that of Walker et al. used judgments of profundity. Finally, participants in the current study were explicitly told to rely on their intuition, and some participants

may have followed this instruction more strictly than others. Despite the differences between studies, there *was* an isolated positive correlation of apophenia with the perceived meaningfulness of CI sentences in the current study, which should be further investigated in future work.

Interestingly, participants' reactions to the two additional sentence types tested, namely sensible sentences (*Your teacher can open the door, but you must enter by yourself*) and nonsense sentences (*Instead of big fish, you can also rarely flip seven potatoes*) did show some clear correlations: People who tended to see more meaning in nonsense tended to see *less* meaning in sensible sentences, and vice versa. Furthermore, this tendency was connected to participants' depth of processing: The reading times of participants who distinguished less between nonsense and sensible sentences were less affected by sentence length, suggesting that they were paying less attention. In light of this findings, it is not clear why bullshit sentences and illusion sentences should be unaffected by the depth-of-processing effect, especially given that semantic illusions have been hypothesized to involve superficial processing (Wason & Reich 1979, Paape et al. 2020, Paape 2021, Leivada 2020). However, there was also no evidence in the data that finding meaning in bullshit and illusion sentences correlates with *deeper* processing, which might be the case if the sentences can be coerced or "repaired" into meaningfulness via additional reasoning steps and/or linguistic operations (Dalton 2016, O'Connor 2015, Wellwood et al. 2018, Zhang et al. 2022). It is possible that depth-of-processing effects are simply more subtle for "semi-meaningful" sentences than for clearly sensible or clearly nonsensical sentences. Thus, further research with larger participant samples, and ideally with more measurements per participant, is clearly needed.

Overall, despite the lack of conclusive results, the present work has highlighted the untapped potential of psycholinguistic investigations into bullshit processing. Even though it has proven challenging to define bullshit in terms of linguistic features (e.g., Cohen 2002), there are some notable tendencies, such as the heavy use of nouns and of abstract rather than concrete words (Turpin et al. 2019, Buekens & Boudry 2015), as well as the heavy use of "genre"-specific jargon (Spicer 2020). Each of these features may make unique contributions to bullshit processing, as well as to individual differences in bullshit receptivity. Applying the entirety of the empirical (psycho)linguistic toolbox to bullshit sentences and investigating connections with other (psycho)linguistic phenomena will undoubtedly lead to many valuable insights in this domain.

Acknowledgments. The study was funded by the University of Potsdam. The author would like to thank the Vasishth Lab members for helpful comments and suggestions.

References

de Almeida, Roberto G. 2018. Composing meaning and thinking. In Gerhard Preyer (ed.), *Beyond semantics and pragmatics*, 201–229. Oxford: Oxford University Press.

Bainbridge, Timothy F, Joshua A Quinlan, Raymond A Mar & Luke D Smillie. 2019. Openness/intellect and susceptibility to pseudo–profound bullshit: A replication and extension. *European Journal of Personality* 33(1). 72–88. https://doi.org/10.1002/per.2176.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

van den Broek, P., C. M. Bohn-Gettler, P. Kendeou, S. Carlson & M. J. White. 2011. When a

reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. Magliano & G. Schraw (eds.), *Text relevance and learning from text*, 123–139. Greenwich, CT: Information Age.

van den Broek, Paul, Robert F Lorch, Tracy Linderholm & Mary Gustafson. 2001. The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition* 29(8). 1081–1087. https://doi.org/10.3758/BF03206376.

Buekens, Filip & Maarten Boudry. 2015. The dark side of the loon. Explaining the temptations of obscurantism. *Theoria* 81(2). 126–142. https://doi.org/10.1111/theo.12047.

Čavojová, Vladimíra, Ivan Brezina & Marek Jurkovič. 2020. Expanding the bullshit research out of pseudo-transcendental domain. *Current Psychology* 41. 827–836. https://doi.org/10.1007/s12144-020-00617-3.

Christianson, Kiel. 2016. When language comprehension goes wrong for the right reasons: Good-enough, underspecified, or shallow language processing. *Quarterly Journal of Experimental Psychology* 69(5). 817–828. https://doi.org/10.1080/17470218.2015.1134603.

Cohen, Gerald A. 2002. Deeper into bullshit. In Lee Overton & Sarah Buss (eds.), *Contours of agency: Essays on themes from Harry Frankfurt*, 321–339. Cambridge, MA: MIT Press.

Cook, Paul & Suzanne Stevenson. 2010. No sentence is too confusing to ignore. In *Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the Common Ground*, 61–69.

Dalton, Craig. 2016. Bullshit for you; transcendence for me. A commentary on "On the reception and detection of pseudo-profound bullshit". *Judgment and Decision Making* 11(1). 121–122.

De Dios-Flores, Iria. 2016. More people have presented in conferences than I have. Comparative illusions: When ungrammaticality goes unnoticed. In Aitor Ibarrola-Armendariz & Jon Ortiz de Urbina Arruabarrena (eds.), *On the Move: Glancing Backwards to Build a Future in English Studies*, 219–228. Bilbao: Universidad de Deusto, Servicio de Publicaciones.

DeYoung, Colin G, Rachael G Grazioplene & Jordan B Peterson. 2012. From madness to genius: The openness/intellect trait domain as a paradoxical simplex. *Journal of Research in Personality* 46(1). 63–78. https://doi.org/10.1016/j.jrp.2011.12.003.

Diamond, Cora. 1981. What nonsense might be. *Philosophy* 56(215). 5–22. https://www.jstor.org/stable/3750713.

Evans, Anthony, Willem Sleegers & Žan Mlakar. 2020. Individual differences in receptivity to scientific bullshit. *Judgment and Decision Making* 15(3). 401–412.

Fortuin, Egbert. 2014. Deconstructing a verbal illusion: The 'No X is too Y to Z' construction and the rhetoric of negation. *Cognitive Linguistics* 25(2). 249–292. https://doi.org/10.1515/cog-2014-0014.

Fowler, Roger. 1969. On the interpretation of 'nonsense strings'. *Journal of Linguistics* 5(1). 75–83. https://www.jstor.org/stable/4175019.

Frankfurt, Harry. 1986. On bullshit. *Raritan Quarterly Review* 6(2). 81–100. Republished as On Bullshit by Princeton University Press, 2005.

Franklin, Michael S, Jonathan Smallwood & Jonathan W Schooler. 2011. Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review* 18(5). 992–997. https://doi.org/10.3758/s13423-011-0109-6.

Gelman, A. & J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*.

Cambridge, UK: Cambridge University Press.
Iacobucci, Serena & Roberta De Cicco. in press. A literature review of bullshit receptivity: Perspectives for an informed policy making against misinformation. *Journal of Behavioral Economics for Policy* 6(1). 23–40.
Kruschke, John. 2014. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. New York: Academic Press.
Leivada, Evelina. 2020. Language processing at its trickiest: Grammatical illusions and heuristics of judgment. *Languages* 5(3). 29. https://doi.org/10.3390/languages5030029.
Lewandowski, Daniel, Dorota Kurowicka & Harry Joe. 2009. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100(9). 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008.
Liddell, Torrin M & John K Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79. 328–348. https://doi.org/10.1016/j.jesp.2018.08.009.
Makowski, Dominique, Mattan S. Ben-Shachar & Daniel Lüdecke. 2019. bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software* 4(40). 1541.
O'Connor, Ellen. 2015. Comparative illusions at the syntax-semantics interface. *Los Angeles, CA: University of Southern California dissertation* .
Paape, Dario. 2021. The role of incremental and superficial processing in the depth charge illusion: Experimental and modeling evidence. 10.31234/osf.io/jp2ma. PsyArXiv preprint. psyarxiv.com/jp2ma.
Paape, Dario, Shravan Vasishth & Titus von der Malsburg. 2020. Quadruplex negatio invertit? The on-line processing of depth charge sentences. *Journal of Semantics* 37(4). 509–555. https://doi.org/10.1093/jos/ffaa009.
Palan, Stefan & Christian Schitter. 2018. Prolific.ac – A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17. 22–27. https://doi.org/10.1016/j.jbef.2017.12.004.
Pennycook, Gordon, James Allan Cheyne, Nathaniel Barr, Derek J Koehler & Jonathan A Fugelsang. 2016. It's still bullshit: Reply to Dalton (2016). *Judgment and Decision Making* 11(1). 123–125.
Pennycook, Gordon, James Allan Cheyne, Nathaniel Barr, Derek J Koehler, Jonathan A Fugelsang et al. 2015. On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making* 10(6). 549–563.
Petrocelli, John V, Haley F Watson & Edward R Hirt. 2020. Self-regulatory aspects of bullshitting and bullshit detection. *Social Psychology* 51(4). 239–253. https://doi.org/10.1027/1864-9335/a000412.
Pinheiro, J. C. & D. M. Bates. 2000. *Mixed-effects models in S and S-PLUS*. New York: Springer.
van Prooijen, Jan-Willem, Karen M Douglas & Clara De Inocencio. 2018. Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural. *European Journal of Social Psychology* 48(3). 320–335. https://doi.org/10.1002/ejsp.2331.
R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for

Statistical Computing Vienna, Austria. https://www.R-project.org/.

Reichle, Erik D, Andrew E Reineberg & Jonathan W Schooler. 2010. Eye movements during mindless reading. *Psychological Science* 21(9). 1300–1310. https://doi.org/10.1177/0956797610378686.

Rule, Alix & David Levine. 2012. International Art English: On the rise, and the space, of the art world press release. *Triple Canopy* 16. 7–30.

Sanford, Anthony J & Patrick Sturt. 2002. Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences* 6(9). 382–386. https://doi.org/10.1016/S1364-6613(02)01958-7.

Schad, Daniel J., Antje Nuthmann & Ralf Engbert. 2012. Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition* 125(2). 179–194. https://doi.org/10.1016/j.cognition.2012.07.004.

Sperber, Dan. 2010. The guru effect. *Review of Philosophy and Psychology* 1(4). 583–592. https://doi.org/10.1007/s13164-010-0025-0.

Spicer, André. 2020. Playing the bullshit game: How empty and misleading communication takes over organizations. *Organization Theory* 1(2). 2631787720929704. https://doi.org/10.1177/2631787720929704.

Stan Development Team. 2022. *Stan modeling language users guide and reference manual*. https://mc-stan.org. Version 2.26.8.

Turpin, Martin Harry, Alexander Walker, Mane Kara-Yakoubian, Nina N Gabert, Jonathan Fugelsang & Jennifer A Stolz. 2019. Bullshit makes the art grow profounder. *Judgment and Decision Making* 14(6). 658–670.

Walker, Alexander, Martin Harry Turpin, Jennifer A Stolz, Jonathan Fugelsang & Derek Koehler. 2019. Finding meaning in the clouds: Illusory pattern perception predicts receptivity to pseudo-profound bullshit. *Judgment and Decision Making* 14(2). 109–119.

Wason, Peter C & Shuli S Reich. 1979. A verbal illusion. *Quarterly Journal of Experimental Psychology* 31(4). 591–597. https://doi.org/10.1080/14640747908400750.

Wellwood, Alexis, Roumyana Pancheva, Valentine Hacquard & Colin Phillips. 2018. The anatomy of a comparative illusion. *Journal of Semantics* 35(3). 543–583. https://doi.org/10.1093/jos/ffy014.

Zhang, Yuhan, Rachel Ryskin & Edward Gibson. 2022. A noisy-channel approach to depth-charge illusions. *Preprint available at SSRN: https://ssrn.com/abstract=4130042* http://dx.doi.org/10.2139/ssrn.4130042.