Tracing Truth Through Conceptual Scaling: Mapping People's Understanding of Abstract Concepts

Lukas S. Huber (lukas.s.huber@unibe.ch)
University of Bern & University of Tübingen

David-Elias Künstle (david-elias.kuenstle@uni-tuebingen.de)
University of Tübingen

Kevin Reuter (kevin.reuter@philos.uzh.ch)
University of Zurich

Abstract

Traditionally, the investigation of *truth* has been anchored in a priori reasoning. Cognitive science deviates from this tradition by adding *empirical* data on how people understand and use concepts. Building on psychophysics and machine learning methods, we introduce *conceptual scaling*, an approach to map people's understanding of abstract concepts. This approach, allows computing participant-specific *conceptual maps* from obtained ordinal comparison data, thereby quantifying perceived similarities among abstract concepts. Using this approach, we investigated individual's alignment with philosophical theories on truth and the predictive capacity of conceptual maps. Obtained results indicated that, while people's understanding of truth is multifaceted and encapsulates notions of coherence and authenticity, alignment is best for the correspondence theory of truth. Furthermore, conceptual maps allowed predicting individual outcomes with an accuracy of ∼70%. This research demonstrates that conceptual scaling offers accurate descriptions of individual's understanding of abstract concepts, behavioral predictions, and quantification of alignment with theoretical perspectives.

Keywords: truth; conceptual scaling; ; concepts; embedding; psychophysics; ordinal comparisons

Introduction

Are we living in a time where truth has become more elusive and subjective, as suggested by d'Ancona (2017), Keyes (2004), and McIntyre (2018)? This prompts questions about what motivates people to deem certain statements as *true* even when they clash with observable evidence. Historically, the philosophical challenge of defining *truth* has persisted as one of the most enduring puzzles. Two theories have received the greatest share of attention. The **correspondence theory** posits that *truth* depends on how well a statement or belief matches up with reality or factual evidence, as argued by Russell (1912) and Wittgenstein (1922). Advocates of the **coherence theory**, including Putnam (1981) and Young (2001), view truth as coherence with a particular set of beliefs. Semantic pluralists like Pedersen & Lynch (2018) and Wright (2005) contend that the term "true" holds multiple interpretations. For example, "it is true that" can imply "reality is such that" in certain situations, while in others it might mean "in line with belief set X".

Previous Empirical Studies on *Truth*

More recently, scholars have started empirically investigating people's ordinary notion of truth. Studies conducted by Barnard & Ulatowski (2013) and Kölbel (2008) have revealed that "true" is understood differently **across domains**—the coherence notion dominates in maths, and the correspondence notion in the sciences like chemistry—suggesting that "true" has at least two senses as argued for by many *truth* pluralists. Interestingly, a parallel observation emerged within the **empirical domain** (Reuter & Brun, 2022). In one of their experiments, participants were confronted with a scenario in which the main character's statement can be perceived as "true" or "false", depending on whether truth is perceived as correspondence or coherence[1]. Participants were divided, with a slight majority leaning towards "true". This finding has been corroborated by Ricciardi & Martin (2022) and was replicated in English, German, and Mandarin by Reuter (2023). These results raise important questions about how contemporary theories of *truth* and the cognitive strategies people use to evaluate *truth* claims should be interpreted. However, these studies have relied on survey-based methods, which has restricted the depth of our understanding regarding people's conceptions of truth.

In this paper, we would like to introduce a methodology—conceptual scaling—that enables the quantification of subjectively perceived relationships among concepts associated with *truth.* Before introducing and applying the proposed method, we aim to identify which concepts closely relate to our understanding of truth. To achieve this, we start our investigation with a preliminary study using a semantic feature production task to reveal these related concepts.

Preliminary Study

While the correspondence theory emphasizes the role of **facts** and alignment with **reality** in determining the truth of a statement, the coherence theory underscores the significance of **justification** and coherence with **reasons** as the core determinants of *truth.* However,

[1] A comparable scenario can be found in our description of Study 2.

the question arises: Do these theories reflect the primary concepts individuals commonly associate with truth? To identify the key attributes that people inherently link with *truth*, we carried out a semantic feature production task. This task is structured to prompt the rapid recall of concept features, thereby revealing the elements most likely to be considered pertinent by individuals when contemplating a particular phenomenon (Barsalou, 1983; Hampton, 1979; Machery, 2017). While there is no standardized protocol for this task, many researchers align with the methodologies outlined by McRae et al. (2005). Typically, participants are prompted to either enumerate features they associate with a concept, c or to specify features deemed characteristic of c.

Methods and Stimuli A gender-balanced sample of 149 participants ($M_{age} = 39.85$) years were asked to list three characteristic features for the terms "truth", "true", or "falsehood".

Results The authors categorized the responses based on their alignment with the correspondence or coherence theory, or their association with *authenticity*.[2] For both the "truth" and "true" categories, the term "honest" emerged as the predominant response, with approximately 70% of participants naming it (see Figure 1). In the "falsehood" category, about 70% of respondents wrote down the word "lie". Notably, only about 40% of those in the "truth" and "true" groups mentioned terms like "fact", "reality", or related words hinting at a correspondence with factual realities. A even smaller proportion referred to terms associated with the coherence theory, such as "reliable", "provable", or "evidence". This task highlights that the most salient feature linked with "truth" and "true" isn't necessarily aligned with the correspondence or coherence theory but rather hinges on the notion of authenticity exemplified by terms such as **honesty** and **transparency**.

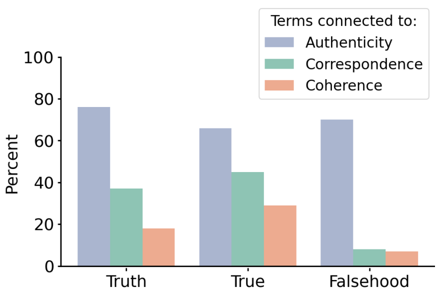

Figure 1: The percentage of responses in a Semantic Feature Production Task for "truth", "true" and "falsehood" that were connected to different theories of *truth*.

[2] Full data set is available in this article's online repository: `https://osf.io/jrfk5/?view_only=56f3460aa3834e9580f70bb3ab132b37`

Discussion The results from this preliminary study highlight "honesty" as the dominant salient feature within our concept of *truth*. Such salience might suggest a complex structure of *truth*, illuminating the complexity of peoples' conceptual understanding. As we will now argue, one way to get a better picture of the conceptual relations of truth with other concepts is through the idea of conceptual spaces, where a quantitative description of similarities and, thus, interrelations among various concepts becomes possible.

Conceptual Spaces

An influential theory in the study of conceptual representations is Peter Gärdenfors' conceptual spaces (e.g., see Gärdenfors, 2004). In Gärdenfors' sense, conceptual spaces are geometric constructs where dimensions correspond to cognitively relevant qualities like color or size, allowing for an intuitive mapping of concepts onto a spatial medium whereby each dimension has a specific, *pre-defined* meaning. This and other related frameworks (e.g., Shepard, 1957; Tversky, 1977, for a review see Roads & Love, 2024) are particularly adept at capturing the nuances of concrete concepts, where each dimension has a direct and interpretable relationship with perceptual experiences. However, regarding abstract concepts such as "truth", the rigidity of pre-defined dimensions may not be adequate. Instead, an unbiased way to describe similarities and interrelations among concepts is required.

We find quantifications of similarity in the field of machine learning as embeddings (Camacho-Collados & Pilehvar, 2018; Goodfellow et al., 2016, Chapter 14) and in psychology as psychophysical scales (Cox & Cox, 2000; Gescheider, 2013a,b). Whereas in machine learning, similarity is typically determined from patterns recognized in large datasets using algorithms, in psychology, similarity is obtained through subjective assessments collected in psychophysical experiments (Wichmann & Jäkel, 2018). However, the same methods also allow for measuring perceived similarities among more abstract stimuli such as images or words (e.g., Arnold, 1971; Henley, 1969; Waraich & Victor, 2023). Traditionally, similarity between pairs of stimuli is assessed by numerical judgments. A more robust procedure—the so-called *triplet comparison task*—however, employs ordinal comparisons to measure relative similarities (Haghiri et al., 2020; Hebart et al., 2020; Lagunas et al., 2019; Roads & Mozer, 2019). By presenting participants with triplets of stimuli and requesting them to evaluate whether stimulus A is more similar to stimulus B or C, this approach provides a more intuitive way to measure perceived similarity among stimuli. From the collected comparisons, machine learning algorithms allow for creating a spatial representation of the stimuli, where distance among points inversely correspond to perceived similarity.

In the present work, we integrate these ideas and methods to study the understanding of abstract concepts. More specifically, we start by presenting words representing abstract concepts and collect similarity judgments from triplet comparisons among those words. Obtained data is then analyzed using an ordinal embedding algorithm as a scaling method. By doing so, subjective similarity judgments are transformed into a structured spatial map uniquely tailored to each participant's understanding. Coordinates within this map accurately reflect the perceived similarities and relations among the concepts used as the basis for the stimuli. We refer to this approach as **conceptual scaling** and to the resulting spatial maps as **conceptual maps**. Conceptual scaling allows us to precisely describe the conceptual understanding of individual participants in a data-driven and empirical way—i.e., without making prior assumptions about potentially uncertain pre-defined dimensions that might otherwise constrain the space a priori. In Study 1, we employ conceptual scaling to investigate the conceptual space of "truth".

Study 1—Tracing Truth

In this study, we run an online experiment to estimate conceptual maps of individual participants. The spatial arrangement of truth-related terms in these 2D maps provides a detailed perspective on individuals' understanding of *truth* and allows investigating the alignment of participants' conceptual understanding of truth with the focal theoretical perspectives on *truth* (correspondence, coherence, authenticity).

Methods

Participants We recruited a gender-balanced sample of 200 participants from the UK and the United States through the online platform Prolific. Out of the 200 participants, a total of $n = 194$ ($M_{age} = 42.87$, $SD_{age} = 13.42$) completed the task.

Stimuli All employed stimuli were written words reflecting concepts. Besides "truth" we used six other words, two for each theoretical perspective. These stimuli were: "reality" & "fact" for correspondence, "judgement" & "reason" for coherence, and "honesty" and "transparency" for authenticity. These stimuli were selected based on a combination of theoretical principles (specifically, words embodying the predominant ideas of the correspondence and coherence theory), empirical evidence from our preliminary study reported above (for instance, "honesty" emerged as the most commonly cited term), and practical considerations—we consciously excluded terms such as "coherence" which might not be commonly used or understood in everyday language.

Triplet Comparison Task Participants engaged in a comparison task where they had to judge similarities among stimuli triplets. A triplet of terms was presented in each trial, and participants had to rate which of the two terms, presented side-by-side, was perceived to be more similar to the term presented above. Their responses were collected using the `left` and `right` arrow keys. Overall, 97 trials were presented with 79 unique triplets and two repetitions of 9 triplets. Unique triplets were randomly selected from all 105 possible triplets ($3\binom{7}{3}$), based on a heuristic commonly used in psychophysical studies (e.g., see Haghiri et al., 2020; Künstle et al., 2022). The repetitions were presented to assess the response consistency among participants.

Procedure From the crowd sourcing platform Prolific, participants were forwarded to a survey on Qualtrics to collect demographic information. Subsequently, participants were routed to an `PsychoPy` (Peirce et al., 2019) experiment hosted online on Pavlovia. Before being presented with the triplet set described above, participants completed two practice sessions: Firstly, they were presented with 10 triplets whereby for each triplet, one target term was clearly more similar to the anchor (e.g., anchor: thunder; target terms: lightning; ketchup). Secondly, participants were presented with 10 triplets featuring uniform similarity relations (e.g., anchor: diamond; target terms: emerald; sapphire). In the main experiment, the 97 trials were split into blocks of 19 triplets.[3]. Between blocks, they had the opportunity for a short break, and after 57 trials, they were forced to pause for two minutes.

Consistency as a Data Quality Indicator

Analysis To determine the reliability of the participants and their conceptual maps, we compute a consistency score, indicating the proportion of repeated triplets in which all three responses were identical. Low consistency may suggest task-related challenges, including inattentiveness, lack of motivation, or misunderstanding of instructions. Theoretically, if participants were responding at random, we would expect a consistency of $(\frac{1}{2})^3 = 12.5\%$

Results Participants consistency scores ranged from 66.66% to 92.59% ($M = 75.54\%, SD = 4.90\%$). This suggests participants' responses were not random but instead exhibited a clear pattern of consistency.

Conceptual Maps

Analysis To estimate individual conceptual maps from participants' responses, we used the conceptual scaling approach described above. Within estimated maps, terms are represented as 2D coordinates whose distances mirror the perceived similarities measured in the triplet comparison task.[4] For example, if a par-

[3] Since 97 is a prime number, the last block consisted of 21 trials.

[4] The choice of a 2-dimensional map is arbitrary and only serves the purpose of simple visualization.

ticipant consistently judge “fact” more similar to “reality” than to “reason”, the conceptual map should reflect this judgement, positioning “fact” closer to “reality” and farther from “reason”. Translating obtained similarity judgments into a precise spatial representation involves a complex optimization task, for which we employed the soft ordinal embedding algorithm (Terada & Luxburg, 2014) implementation in the Python toolbox `cblearn`.[5] In Figure 2 we show a visualization of an individual conceptual map obtained through conceptual scaling (maps of all participants are available in the online repository).

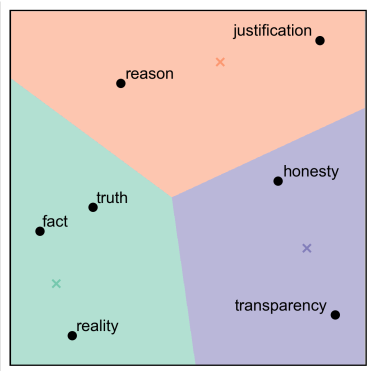

Figure 2: Conceptual map of a particular participant where each truth-related term was assigned to a theory of truth. The colored regions indicate a clear separation of the map by the mean distance to the centroid of each theory (colored x-markers). Since “truth” lies in the region of “fact” and “reality”, this suggests that the participant aligns with the correspondence theory.

Theory-driven Clusters Assignments

Analysis To investigate how well the theoretical similarity of truth-related terms is mirrored in participants' conceptual maps, we assign each term—except the term “truth” itself–to one of the theories: “reality” & “fact” to correspondence, “justification” & “reason” to coherence, and “honesty” and “transparency” to authenticity. Agreement between participants' conceptual maps and this *theory-driven cluster assignment* is quantified by a metric called Silhouette Score (Rousseeuw, 1987).[6] This standard cluster metric evaluates the average degree of similarity (-1 to 1) of each data point to its assigned cluster compared to other clusters. A positive score indicates cohesion within the cluster and separation from other clusters, indicating a clear distinction between the terms assigned to each theory, as in the example in Figure 2. A score around and below 0 suggests overlapping clusters with no clear distinction, which could be interpreted as a missing agreement between participants' conceptual maps and the theory-driven cluster assignment.

Results While lacking agreement could be explained by a different, multi-faceted understanding of truth-related terms, it could also indicate poor data quality. Therefore, in Figure 3, we plot the Silhouette scores as a function of consistency. The reasonably high consistency and the lack of correlation imply that participants with low Silhouette Scores were not providing noisy data but perceived truth-related terms differently from our theory-driven cluster assignment. Most participants, however, show positive scores and an organization of the maps, as expected by the theory-driven assignments. This observation was confirmed by a Wilcoxon signed rank test , indicating that the median Silhouette Score was significantly above zero ($Mdn = 0.13, W = 3563, p < .001$).

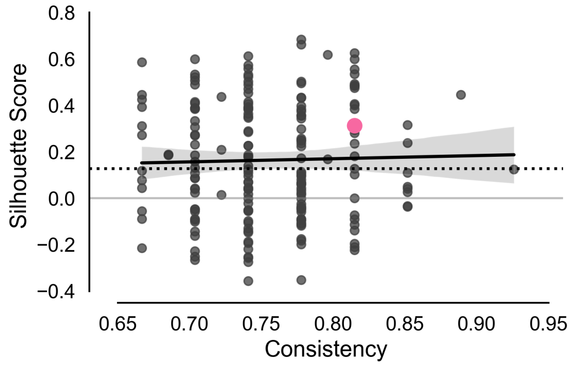

Figure 3: Silhouette Scores as a function of response consistency. Silhouette Scores indicate the level of agreement with theory-driven cluster assignment (higher is better). The dotted horizontal line marks the median Silhouette Score. Consistency indicates how reliably participants responded to the same stimuli across repeated presentations. A simple regression fit indicates no significant association between Consistency and Silhouette score (solid black line; surrounding shaded area indicates the 95% confidence interval). The colored data point highlights the specific participant whose conceptual map is shown in Figure 2.

Theory-driven Conceptual Maps

Analysis To quantify the alignment of participants' conceptual understanding with the focal theoretical perspectives on *truth*, we generated three synthetic sets of triplet responses. Each set represents the hypothetical judgments of a prototypical advocate for one of the focal theories. We apply conceptual scaling to these synthetic responses to get a *theory-driven conceptual map* for each theory. Next, we quantify the alignment between each participant's map and these theory-driven maps in terms of Procrustes disparity (Gower, 1975), which is the sum of point-wise Euclidean distances after optimal similarity transformations (translation, rotation, scaling, flipping).[7] A visualisation of this analysis can be found in Figure 4.

---

[5] An extensive introduction to ordinal embedding algorithms in the context of psychological experiments can be found in Haghiri et al. (2020).

[6] As implemented in the `scikit-learn` library (Pedregosa et al., 2011, version 1.2.2) for Python.

[7] Procrustes analysis was used as implemented in the `scipy` library (Virtanen et al., 2020) for Python.

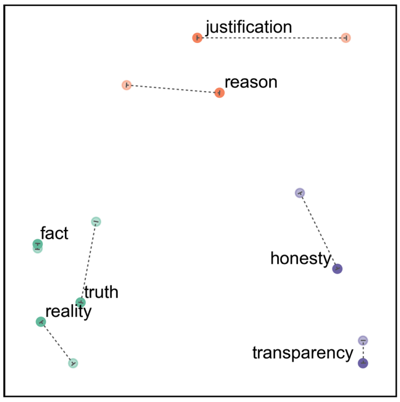

Figure 4: Measuring the alignment between a participant's conceptual map (slightly transparent markers) and a theory-driven conceptual map (here correspondence) in terms of Procrustes disparity, indicated by the dashed lines reflecting point-wise distances. Maps for the other theories (coherence and authenticity) can be found in this article's online repository.

Results Participants' alignment with theory-driven maps varied significantly among theoretical perspectives. Procrustes disparity were lowest for correspondence ($Mdn = 0.50$), followed by authenticity ($Mdn = 0.58$), and coherence ($Mdn = 0.62$). Normality tests (Shapiro-Wilk) indicated non-normal distributions. To evaluate whether alignment with different theories differs significantly, we employed a Friedman test (a non-parametric alternative to the one-way ANOVA with repeated measures), which revealed significant differences ($\chi^2 = 78.48, p < .0001$), with posthoc tests confirming differences between all pairs of perspectives. These findings indicate that participants' conceptual alignment was strongest with the correspondence theory, followed by authenticity, and weakest with the coherence theory.

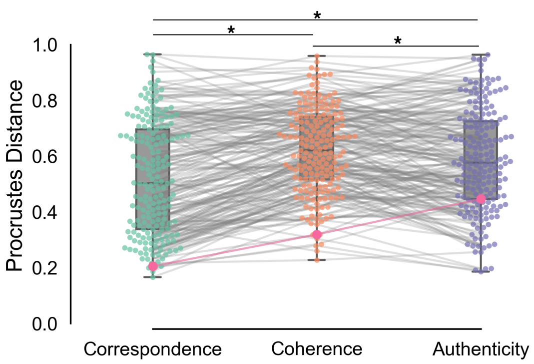

Figure 5: Distribution of Procrustes disparities across all theory-driven conceptual maps. Lower values indicate better alignment. Alignment is best for the correspondence theory, followed by authenticity, and lowest for Coherence. Data points from the same participants are connected by grey lines. Data from the participant whose conceptual map is shown in Figure 2 is highlighted in pink. Asterisks mark significant differences.

Discussion

In this study, we applied conceptual scaling to map people's understanding of *truth*. From the data collected in an online experiment, we estimated a conceptual map of truth-related terms for each of 193 participants. Besides quantitatively describing individuals' conceptual understanding, these maps were employed to investigate alignment with various theories of *truth* in terms of theory-driven cluster assignments and theory-driven maps.

While the maps of most participants are segmented in accordance with theory-driven cluster assignments, there is a relevant part of participants with a deviating spatial arrangement of truth-related terms. This variance in spatial arrangement is not correlated with the participants' response consistency and, thus, most likely not due to task-related artifacts such as a lack of motivation. We also find variability in the alignment with theory-driven conceptual maps: Many participants align with one of the theories, with a slight but significant preference for correspondence, followed by authenticity and coherence. Our results resonate with those reported by Reuter & Brun (2022), yet extend beyond by offering a nuanced, continuous assessment of conceptual understanding and alignment with theoretical perspectives, rather than a mere categorical measurement.

Study 2—Predicting Vignette Responses

While study 1 reveals the potential of conceptual scaling in mapping the understanding of abstract concepts, Study 2 seeks to determine if these conceptual maps also correlate with more contextualized measurements of people's use of concepts. To do so, we investigated whether individual responses in a scenario-based tasks, can be predicted using distance relations obtained form conceptual maps. More specifically, predictions were obtained by calculating the Euclidean distance from "truth" to the centroids of the correspondence ("fact" & "reality") and authenticity ("honesty" & "justification") cluster (for an intuitive understanding see Figure 2).[8] Using this measure of conceptual proximity, we predicted whether participants' responses would either favor the correspondence or authenticity theory.

Methods

Participants After a **delay of three months**, we re-engaged the same participants from Study 1 through the platform Prolific. Out of the 193 participants, 129 completed the task within 48 hours. We reasoned that valid predictions require participants' conceptual maps to show similarity between terms linked to each truth theory, reflecting the structured organization of truth-related terms. Therefore, we limited further analysis to

8 Coherence theory was omitted because responses favoring coherence would collapse with responses favoring authenticity.

participants who achieved a Silhouette Scores above the median in Study 1 (see dashed line in Figure 3). After these exclusions, the sample consisted of $n = 58$ participants ($M_{age} = 48.27$, $SD_{age} = 13.42$; gender distribution (female/male/other): 41.07%/58.93%/0.00%).

Stimuli In line with the experimental paradigm of (Reuter & Brun, 2022), participants read the following vignette:

Maria and Peter are students and meet up for a late dinner. Peter asks Maria whether Tom is at the party that they intend to go to after dinner. Maria answers that Tom is at the party. After all, Tom had told her that he would be at the party. When they arrive at the party, it turns out that Tom had changed his plans, and is not at the party.

Participants were then asked: Was Maria's answer true or false? Participants choose from the two options "true" and "false".

Hypothesis For participants in whose conceptual maps the distance from "truth" to the centroid of the correspondence cluster is shorter than to the centroid of the authenticity cluster, we predict that they will label Maria's answer as "false", and the reverse for those closer to the authenticity cluster.

Results

55.17% of participants responded that Maria's answer was "false" (agreeing with correspondence theory) and 44.83% that her answer was "true" (agreeing with a authenticity notion of "truth"). We were able to correctly predict 67.50% of the correspondence and 72.22% of authenticity responses, resulting in a mean prediction accuracy of 69.86%. A chi-square test confirmed the significance of the correlation between predicted and actual responses, $\chi^2(1, N = 64) = 6.39, p = .011$). With a Cramér's V of 0.31, this result denotes a moderate association, indicating that the distances within conceptual maps are effective predictors of how participants judge truth-related scenarios.

Discussion

These results suggest that conceptual maps can meaningfully predict behavioral outcomes in tasks closer to everyday life, affirming their ability to accurately reflect individual participants' conceptual understanding. The notable predictive accuracy of nearly 70%, maintained even after a 3-month interval between the estimation of the conceptual maps and the measurement of the outcome, underscores the robustness and stability of conceptual maps constructed through conceptual scaling.

General Discussion

Summary of Studies and Results In Study 1, we employed conceptual scaling to measure people's understanding of *truth*. Estimated conceptual maps allowed us to precisely quantify participants' alignment with different theoretical notions about *truth* and revealed a multifaceted and heterogeneous understanding. While people's understanding of *truth* encapsulates notions of coherence and authenticity, alignment is best for the correspondence theory of truth. In Study 2, we re-engaged the same participants and found that obtained conceptual maps proved effective in predicting response behavior in a more contextually embedded task.

Philosophical Implications Most philosophers typically advocate for one of three principal theories of truth: the correspondence theory, the coherence theory, or a domain-specific pluralistic theory. Our research findings, however, present a significant challenge to these established divisions in two key ways. First, while *honesty* is traditionally associated primarily with truthfulness, our studies reveal that concepts like *honesty* and *authenticity* are also crucial in shaping our understanding of truth. This intertwining of ideas has been largely overlooked in previous discourse. Second, our results bolster the argument for a comprehensive pluralism in conceptions of truth, extending across various domains and within the empirical realm itself.

Applications and Limitations Conceptual scaling emerges as a potent tool to map people's understanding of abstract concepts—offering precise quantification, the possibility to measure alignment, and long-term predictability. The effectiveness in measuring participant-specific conceptual understanding with minimal contextual dependency may be suitable for controlled experiments where other methods, such as corpus analysis or context-dependent methods, reach their limits. Our current approach also has some limitations. Future research should aim to extend the validation of conceptual scaling to a broader range of concepts to solidify the method's applicability. Additionally, while the dimensions of the conceptual maps offer a spatial representation of conceptual relationships, they are mathematical constructs from the embedding algorithm, and their direct semantic interpretability remains difficult.

Conclusion

Drawing inspiration from various fields, we have presented a novel method—conceptual scaling—that enables researchers to track and gain a deeper insight into abstract concepts. This methodology was effectively applied to the concept of truth, serving as a proof of concept and showcasing its potential. We hope our approach proves helpful to advance the understanding and mapping of abstract concepts within individual minds.

Acknowledgments

We would like to thank Felix A. Wichmann for valuable discussions and insightful suggestions that contributed to the development of this work. Thanks also to the audience of the Zurich XPhi Lab Meeting for helpful discussions on the topic. Lukas S Huber was funded by the Swiss National Science Foundation (grant number: 214659). The research of Kevin Reuter was also supported by the Swiss National Science Foundation (grant number PCEFP1_181082). The research of David-Elias Künstle is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC number 2064/1 – Project number 390727645 and the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting David-Elias Künstle.

References

Arnold, J. B. (1971). A multidimensional scaling study of semantic distance. *Journal of Experimental Psychology*, *90*(2), 349.

Barnard, R., & Ulatowski, J. (2013). Truth, correspondence, and gender. *Review of Philosophy and Psychology*, *4*, 621–638.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *10*, 82–93.

Camacho-Collados, J., & Pilehvar, M. T. (2018). From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, *63*, 743–788.

Cox, T., & Cox, M. (2000). *Multidimensional Scaling (2nd ed.)*. Chapman and Hall/CRC.

d'Ancona, M. (2017). *Post truth: The new war on truth and how to fight back.* London: Random House.

Gescheider, G. A. (2013a). The Measurement of Sensory Attributes and Discrimination Scales. In *Psychophysics: The Fundamentals* (pp. 183–206). Taylor & Francis.

Gescheider, G. A. (2013b). Partition Scales. In *Psychophysics: The Fundamentals* (pp. 183–206). Taylor & Francis.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, *40*, 33–51.

Gärdenfors, P. (2004). *Conceptual spaces: The geometry of thought.* MIT press.

Haghiri, S., Wichmann, F. A., & von Luxburg, U. (2020, September). Estimation of perceptual scales using ordinal embedding. *Journal of Vision*, *20*(9), 14. doi: 10.1167/jov.20.9.14

Hampton, J. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning & Verbal Behavior*, *18*, 441-461.

Hebart, M. N., Zheng, C. Y., Pereira, F., & Baker, C. I. (2020, November). Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nature Human Behaviour*, *4*(11), 1173–1185. doi: 10.1038/s41562-020-00951-3

Henley, N. M. (1969, April). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 176–184. doi: 10.1016/S0022-5371(69)80058-7

Keyes, R. (2004). *The post-truth era: Dishonesty and deception in contemporary life.* New York: Macmillan.

Künstle, D.-E., von Luxburg, U., & Wichmann, F. A. (2022). Estimating the perceived dimension of psychophysical stimuli using triplet accuracy and hypothesis testing. *Journal of Vision*, *22*(13), 5–5.

Kölbel, M. (2008). "true" as ambiguous. *Philosophy and Phenomenological Research*, *77*(2), 359–384.

Lagunas, M., Malpica, S., Serrano, A., Garces, E., Gutierrez, D., & Masia, B. (2019, July). A Similarity Measure for Material Appearance. *ACM Transactions on Graphics*, *38*(4), 1–12. doi: 10.1145/3306346.3323036

Machery, E. (2017). *Philosophy within its proper bounds.* Oxford: Oxford University Press.

McIntyre, L. (2018). *Post-truth.* Cambridge, MA: MIT Press.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559.

Pedersen, N. J. L. L., & Lynch, M. P. (2018). Truth pluralism. In M. Glanzberg (Ed.), *The Oxford handbook of truth* (pp. 543–575). Oxford: Oxford University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior research methods*, *51*, 195–203.

Putnam, H. (1981). *Reason, truth and history.* Cambridge: Cambridge University Press.

Reuter, K. (2023). The ambiguity of 'true' in English, German and Chinese.
(Unpublished Manuscript)

Reuter, K., & Brun, G. (2022). Empirical studies on truth and the project of reengineering truth. *Pacific Philosophical Quarterly*, *103*(3), 493-517.

Ricciardi, G., & Martin, J. (2022). Accounting for variability in the truth-evaluation of bare epistemic possibility statement. *Proceedings of the Linguistic Society of America*, *7*(1), 5249.

Roads, B. D., & Love, B. C. (2024). Modeling similarity and psychological space. *Annual Review of Psychology*, *75*(1), 215-240. (PMID: 37562499) doi: 10.1146/annurev-psych-040323-115131

Roads, B. D., & Mozer, M. C. (2019, October). Obtaining psychological embeddings through joint kernel and metric learning. *Behavior Research Methods*, *51*(5), 2180–2193. doi: 10.3758/s13428-019-01285-3

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, *20*, 53–65.

Russell, B. (1912). *The problems of philosophy.* Oxford University Press.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22*(4), 325–345.

Terada, Y., & Luxburg, U. (2014). Local ordinal embedding. In *International conference on machine learning* (pp. 847–855).

Tversky, A. (1977). Features of similarity. *Psychological review*, *84*(4), 327.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: 10.1038/s41592-019-0686-2

Waraich, S. A., & Victor, J. D. (2023). The geometry of low- and high-level perceptual spaces. *Journal of Neuroscience*. doi: 10.1523/JNEUROSCI.1460-23.2023

Wichmann, F. A., & Jäkel, F. (2018, March). Methods in Psychophysics. In J. T. Wixted (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (pp. 1–42). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9781119170174.epcn507

Wittgenstein, L. (1922). *Tractatus logico-philosophicus.* London & Henley: Routledge & Kegan Paul. (English translation by D. F. Pears & B. F. McGuinness)

Wright, C. (2005). On the functionalization of pluralist approaches to truth. *Synthese*, *145*(1), 1–128.

Young, J. (2001). A defence of the coherence theory of truth. *The Journal of Philosophical Research*, *26*, 89–101.