

Using comparison judgments to study representations

Steven Langsford
The University of Adelaide

October 31, 2017

Abstract

Three projects are presented, all using comparison data to investigate representations. Processes of comparison are the focus here because of the strong links they create between the abstract representations much of cognition aims to study and unambiguous choice outcomes. The superficial similarities between these projects, that they all use browser based studies to reach relatively large numbers of people, and apply quantitative models to summarize and interpret the results, derive from two things: a common set of concerns with representation structure, and the use of comparison tasks to contrive situations where different representations predict different task behaviors. These basic ideas are applied across different domains to address current questions of representation and measurement in similarity and language.

The first section compares two prominent theories of similarity judgment, transformational similarity and structural alignment, across three studies. The first of these constructs triad stimuli such that the two approaches make opposite predictions, the second measures similarity using an alternative measure of same-different discrimination speed, and a third applies both tasks to a common set of stimuli to clearly resolve their similarities and differences. The results show evidence of a misspecification in the APPLY rule of the transformational account current for geometric shapes, and also show that while same-different discrimination and deliberative comparison measures of similarity judgment are largely consistent, there are differences which appear to arise due to the different time constraints of the two tasks.

The second section investigates a paradigm for testing the impact of transformation learning on similarity and categorization judgments. In this paradigm, a common set of test items follows two different training conditions, such that no test item is present in any training, and the status of each test item as a match, near match, or non-match to the training varies by condition. Responses to identical test items are compared across training conditions to expose the impact of transformation training on similarity and categorization judgment. Across multiple iterations of this basic design I show that the transformations are learned, and that transformation learning does impact similarity and categorization judgment. Change in similarity and categorization ratings due to training are largest in the easiest training conditions where transformations are presented explicitly to participants during training, and less pronounced when transformations are presented implicitly. Some generalization of learning is shown across related transformations, suggesting some similarity structure among transformations.

The third section moves into empirical studies of syntax, comparing different ways of measuring sentence acceptability, the degree to which a sentence appears well-formed to a speaker of that language. This is related to the similarity work in the first and second sections through its use of Thurstonian modeling for structure discovery, which is capable of inferring acceptability scores for each sentence while also avoiding the need to present a rating scale of any kind to participants. This study complements existing work on the Type 1 and Type 2 error rates of the most common measurement techniques with its investigation of within and between participant test-retest reliability. The Likert task is found to be particularly effective. The results presented here show it has particularly good reliability properties and help empirically validate the common practice of interpreting averaged Likert ratings as a fine-grained measure of gradient acceptability.

Dedication

To the Mechanical Turk community. I know the ‘short, fun psychology study’ HITs weren’t always short and fun. But they were all studies: here are the results.

Almost all Mechanical Turk studies implicitly assume that the Mechanical Turk community is reasonably representative of humanity. Working with the Turker community has convinced me that humanity is very smart, very impatient, almost never malicious, and overwhelmingly willing give a good faith effort in doing a task so long as the researcher has made a good faith effort to make the task clear. There’s no particular reason these things had to be true facts about humanity. But they are true, and as a result, all the work presented here was possible. Thank you all.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed:

18/8/2017

Acknowledgements

Science is about people, and the ones at the Adelaide computational cognitive science lab have been great people to try and do science with.

My supervisors for this work, Dani Navarro, Amy Perfors, and Andrew Hendrickson, can't be thanked enough. If this project was a car, they kept an eye on the nuts and bolts to stop the wheels falling off, patiently taught me to drive as I bunny-hopped around thrashing the gears, and gave me a road map with a handy arrow on it marked "You are here." There's no better student experience than that.

Contents

I	General Introduction	8
1	Comparison and representation	9
II	Similarity representation	12
2	Contrasting accounts of similarity	13
2.1	Studying similarity	13
2.1.1	Two approaches to structure	15
2.1.2	When two approaches is one too many	16
3	Transformation or alignment?	18
3.1	Introduction	21
3.1.1	Similarity as structural alignment	22
3.1.2	Similarity as transformation	23
3.2	Experiment 1	23
3.2.1	Method	24
3.2.2	Results	27
3.2.3	Discussion	27
3.3	Experiment 2	28
3.3.1	Method	29
3.3.2	Results	29
3.3.3	Discussion	30
3.4	Experiment 3	31
3.4.1	Method	32
3.4.2	Results	35
3.4.3	Discussion	38
3.5	General Discussion	39
4	Which transformations?	42
5	Transformation learning	45
5.1	Introduction	48
5.2	Experiment 1	48
5.2.1	Method	49
5.2.2	Results	51
5.2.3	Conclusion	52
5.3	Experiment 2	52
5.3.1	Method	52
5.3.2	Results	53
5.4	General Discussion	54
III	Linguistic representation	55
6	From similarity to sentences	56
6.1	The structure of acceptability	56

6.2	Testing the Thurstonian model	58
6.2.1	Simulation tests	58
6.2.2	Interpreting distances on the scale	59
6.3	Reliability studies for sentence acceptability measures	62
7	The reliability of acceptability	65
7.1	Introduction	68
7.1.1	The measures	70
7.1.2	Measure evaluation	71
7.2	Method	72
7.2.1	Sentences	72
7.2.2	Measures	72
7.2.3	General procedure	75
7.2.4	Task-specific procedures	76
7.3	Results	78
7.3.1	Global measures	79
7.3.2	Decision measures	83
7.3.3	Agreement between measures	87
7.4	Summary and Conclusions	88
IV	Conclusions	92
8	Summary and conclusions	93
	References	98
V	Appendices	107
A	Sentence stimuli	108
A.1	Attention check questions	108
A.2	Instruction quiz	108
A.3	Sentence stimuli	108
A.4	Conflict sentences	113
B	Model implementation	114

“Comparison is the death of joy” : Mark Twain

“We can gain intense pleasure only from the contrast”: Freud

“What’s the difference between a zippo and a hippo?”

“One of them is a little lighter.” : Masai Graham.