Original Article

Sociological Methods & Research
1-23

DOI: 10.1177/0049124120914937
journals.sagepub.com/home/smr
SAGE

Q-SpAM: How to Efficiently Measure Similarity in Online Research

Alex Koch[1,2], Felix Speckmann[1] and Christian Unkelbach[1]

Abstract

Measuring the similarity of stimuli is of great interest to a variety of social scientists. Spatial arrangement by dragging and dropping "more similar" targets closer together on the computer screen is a precise and efficient method to measure stimulus similarity. We present Qualtrics-spatial arrangement method (Q-SpAM), a feature-rich and user-friendly online version of spatial arrangement. Combined with crowdsourcing platforms, Q-SpAM provides fast and affordable access to similarity data even for large stimulus sets. Participants may spatially arrange up to 100 words or images, randomly selected targets, self-selected targets, self-generated targets, and targets self-marked in different colors. These and other Q-SpAM features can be combined. We exemplify how to collect, process, and visualize similarity data with Q-SpAM and provide R and Excel scripts to do so. We then illustrate Q-SpAM's versatility for social science, concluding that Q-SpAM is a reliable and valid method to measure the similarity of lots of stimuli with little effort.

[1] Center for Social and Economic Behavior (C-SEB), University of Cologne, Koln, Germany
[2] The University of Chicago Booth School of Business, IL, USA

Corresponding Author:
Alex Koch, The University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637, USA.
Email: alex.koch@chicagobooth.edu

Keywords

similarity, measurement, spatial arrangement method, Qualtrics, Q-SpAM

Similarity is a central and useful construct in the social sciences. As James (1890) stated: “sameness is [. . .] the backbone of our thinking” (p. 459). Although there are theoretical debates on how people compute and represent similarity (Gentner and Markman 1997; Griffiths, Steyvers, and Tenenbaum 2007; Jones and Mewhort 2007; Krumhansl 1978; Medin, Goldstone, and Gentner 1993; Tversky 1977; for a review, see Goldstone and Son 2005), similarity predicts how people think and act. For example, targets similar to templates and dissimilar to distractors are spotted easier (Godwin, Hout, and Menneer 2014; Hout and Goldinger 2015). Prime-target similarity facilitates classification in evaluative priming (Unkelbach et al. 2008). Targets similar to prototypes are categorized more accurately (Rips 1989; Smith and Sloman 1994). Generalizations between similar targets are more likely (Gräf and Unkelbach 2016; Shepard 1987). Distractor-target similarity interferes with recognition but facilitates recall (Alves et al. 2015; Roediger and McDermott 1995). And self-target similarity breeds attraction (Klohnen and Lou 2003), liking (Alves 2018; Brandt 2017; Byrne 1969), compliance to requests (Burger et al. 2004), emotional support (Suitor, Keeton, and Pillemer 1995), trust (Sofer et al. 2015), and even economic cooperation under risk (DeBruine 2002; Fischer 2009).

Researching similarity’s cognitive and behavioral effects and/or predicting targets’ recognition, liking, and so on, from their similarity requires a method to measure it. One direct measure is pairwise judgment on a scale ranging from “similar” to “dissimilar.” However, similarity is highly context dependent. Two targets (e.g., dog and cat) may be judged as more or less similar in the presence of a third target. Given a puppy, dog and cat are dissimilar, but given an axolotl, they are similar (Goldstone, Medin, and Halberstadt 1997). To keep context constant, researchers typically instruct participants to judge the similarity of all $N \times (N - 1)/2$ unique pairs in a relevant set (e.g., 30 targets require 435 pairwise judgments). Pairwise judgment is rarely used to measure the similarity of larger target sets (e.g., eight acquaintances, Alves, Koch, and Unkelbach 2016; 12 brands, Bijmolt et al. 1998) and, in our experience, becomes cumbersome for participants if target sets exceed about 20 targets (i.e., 190 pairwise judgments; see Unkelbach et al. 2008, Study 3).

A more efficient measure is sorting similar targets into the same pile and dissimilar targets into different piles (pairs ending up in the same pile and different piles are coded as 0 ["similar"] and 1 ["dissimilar"], respectively). Pile sorting is more efficient than pairwise judgment (e.g., 135 emotions, Shaver et al. 1987; 220 groups, Pattyn, Rosseel, and Van Hiel 2013) because sorting a target into a pile simultaneously judges the similarity between that target and all other targets already in a pile. However, pile sorting is a binary measure, and thus, its efficiency comes at the cost of precision.

Spatial Arrangement: A Novel Similarity Measure Both Precise and Efficient

A both efficient and precise measure is dragging and dropping "more similar" targets closer together and "more dissimilar" targets further apart on a surface (Goldstone 1994). Spatial arrangement (aka projective mapping,[1] Risvik et al. 1994) is efficient (e.g., 48 states, Koch et al. 2018; 50 occupations, Imhoff, Koch, and Flade 2018) because moving a target simultaneously adjusts the proximity (i.e., judges the similarity) between that target and all other targets already on the surface. The efficiency advantage over sequential pairwise judgment is already the case given a few targets; given many targets, spatial arrangement is substantially more efficient. Spatial arrangement is also precise because a surface provides lots of measuring points between maximum similarity (same spot) and maximum dissimilarity (e.g., opposite ends). It is true that this precision advantage over binary pile sorting comes at the cost of being limited to two orthogonal dimensions (i.e., a surface is two-dimensional, Verheyen et al. 2016). However, participants can imperfectly use more than two orthogonal dimensions to spatially arrange targets on a surface (Hout and Goldinger 2016), and spatial arrangement of multiple target subsets on consecutive surfaces enables perfect usage of more than two orthogonal dimensions (Kriegeskorte and Mur 2012). Thus, we believe the precision advantage of spatial arrangement over pile sorting outweighs its complexity disadvantage at least for targets that most participants mentally organize along a few dimensions only.

Hout, Goldinger, and Ferguson (2013; see also Goldstone 1994; Koch, Aleves et al. 2016; Kriegeskorte and Mur 2012; Risvik et al. 1994) showed that for various visual and verbal targets, spatial arrangement correlated strongly with pairwise judgment, and Koch and colleagues (2018) showed that for various verbal targets, spatial arrangement correlated strongly with pile sorting. Koch, Aleves et al. (2016) confirmed the concurrent validity of spatial arrangement, too, and also showed its predictive validity: Both spatial arrangement and pairwise judgment similarity correlated moderately with classification speed and recognition accuracy. In sum, spatial arrangement is an efficient, precise, and valid method to measure similarity.

Given high efficiency at high precision, spatial arrangement has become a more and more popular similarity measure. On May 03, 2019, Google Scholar indicated 126, 43, 36, and 156 citations of the arguably four most extensive validations of spatial arrangement by Goldstone (1994), Hout, Goldinger, and Ferguson (2013), Koch, Aleves et al. (2016), and Risvik and colleagues (1994), respectively. We believe it would be even more popular if it could be integrated into streams of research in a more automated way. The least automated approach is spatial arrangement on a sheet of paper followed by manual (for tea, Moelich et al. 2017; for wine, Perrin et al. 2008) or photo-based, semiautomated recording of proximity between targets (for animals, Montez, Thompson, and Kello 2015; for plants and animals, Unger et al. 2016).

Hout, Goldinger, and Ferguson (2013; see http://www.michaelhout.com/) provided a more automated approach with a freeware that records proximity between targets that have been spatially arranged on the computer screen. They provided E-Prime code and Java apps for running screen-based spatial arrangement studies with visual or verbal targets. For example, with this freeware, Hout, Goldinger, and Brady (2014) provided similarity data for 240 object categories (e.g., teddies and butterflies; 16 or 17 photographed targets per object category), and Horst and Hout (2016) provided similarity data for 64 photoshopped, novel objects. The targets in these databases will keep stimulating similarity research. For similarity within populous categories (i.e., 20+ targets), similarity between categories, similarity of verbal targets, and for predicting cognition and behavior from similarity within participants, however, new data will have to be collected.

For research projects that require similarity data from a large and diverse participant sample, we argue that online spatial arrangement is a promising venue. Lê, Brard, and Lê (2017) provided Holos, a free Android app with which participants can spatially arrange target words, images, sounds, or videos on their tablet. Holos saves each target's path and final destination on a freely accessible server on which researchers can share their study templates. After creating a Holos account and setting up a study including uploading targets (http://napping.agrocampus-ouest.fr/), researchers invite participants via e-mail. After downloading (same link) and installing Holos on their tablet and entering their name, e-mail, and the study's ID included in their invitation, participants use their fingers to spatially arrange the targets of interest. The efficiency advantage of this automated online approach is that from anywhere with internet connection, hundreds of people can take part in the same spatial arrangement study at the same time (i.e., more than in any lab).

However, researchers need participants' e-mail addresses, participants need a tablet and need to download and install an app, and so far, their data are tied to their e-mail address and name and thus identifiable, limiting the scope of Holos to nonanonymous participants on e-mail lists and regulations in which identifiability of data is not an issue. Given novel legislations, particularly the European Union's General Data Protection Regulation (http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:ST_5419_2016_INIT), anonymity and nonidentifiability are critical issues on all online crowdsourcing platforms (e.g., Mechanical Turk and Prolific Academic; Buhrmester, Kwang, and Gosling 2011; Peer et al. 2017). Thus, combining online spatial arrangement with online crowdsourcing platforms by ensuring participants' anonymity and nonidentifiability would greatly increase the possibility to efficiently collect interstimulus similarity data for large sets of stimuli form a diverse population. The free JavaScript code (embedded in Qualtrics) presented in this article provides this possibility.

Online Spatial Arrangement Powered by Qualtrics (Q-SpAM)

As of 2018, Qualtrics is arguably the global leader in browser-based survey software used by universities (https://www.qualtrics.com/education/higher-education/). Typically, researchers set up a study on Qualtrics' website and post a link to the study alongside with information about its type/topic, duration, payment (typically U.S.$1–2 for 10 minutes), and available slots (i.e., *N*) on an online crowdsourcing platform. Anonymous participants follow the link, complete the Qualtrics study in their browser (i.e., with their desktop, laptop, tablet, or smartphone), and get redirected to the platform and paid[2] what was specified in the study's advertisement. Combining platforms' rapid and affordable access to participants with the payware Qualtrics is highly popular among researchers because it has many user-friendly functions for designing and sharing studies and results. Importantly, by embedding JavaScript code, it is possible to add novel functions. We integrated the spatial arrangement method into Qualtrics (Qualtrics-spatial arrangement method; henceforth Q-SpAM).

In the following, we outline how to set up and customize Q-SpAM. The Online Supplementary Material (which can be found at http://smr.sagepub.com/supplemental/) provides a more detailed manual. The goal here is to provide a sense of what can be done with Q-SpAM. Under separate headings, we explain that participants can spatially arrange up to 100 words or images, randomly selected targets, self-selected targets, self-generated targets, and targets self-marked in different colors; that the background of the spatial arrangement slide can be blank or coordinate axes or circles; and that these and other Q-SpAM features and standard Qualtrics features can be combined. As an illustrative example, we use a set of influential leaders as target stimuli for similarity assessment.

Default Settings of Q-SpAM

One needs a Qualtrics account to implement Q-SpAM. After logging into Qualtrics, one must load "Q SpAM.qsf" (see Open Science Foundation [OSF] website https://osf.io/zkfvh/) from its hard disk location. After loading this file, one can click on project "Q SpAM" and then "preview survey," which provides a preview of what participants will see when doing the survey. Following an initial screen slide, participants will read the spatial arrangement instructions shown in Figure 1A. Next, 20 targets (in our example, influential leaders) will appear in two columns and 10 rows in the center of a blank slide (it extends well beyond as shown in Figure 1B). Figure 1C shows the instructions as repeated at the bottom of this slide throughout spatial arrangement. Figure 1D shows one of many possible solutions (female and male targets on the left and right side, respectively).

Changing Identity and Number of Targets

One may present other targets and more or less than 20 targets. If you present less than the number of targets in a given stimulus set (i.e., the stimulus pool), each participant will spatially arrange a random sample of targets drawn from that pool. The Q-SpAM default settings present a random sample of 20 of 62 targets. There are 10 pools, so you can combine up to 10 random samples of different size. For example, each participant may spatially arrange 40 targets: 30 and 10 targets randomly drawn from a first and second pool, respectively. Across pools 1–10, the maximum possible number of targets is 100.

Changing Appearance and Initial Location of Targets

One may change the pixel width, height, border size, and horizontal and vertical spacing of target boxes both not yet and already spatially rearranged by a participant (see Figures 1B and 1D). One may also change the boxes'

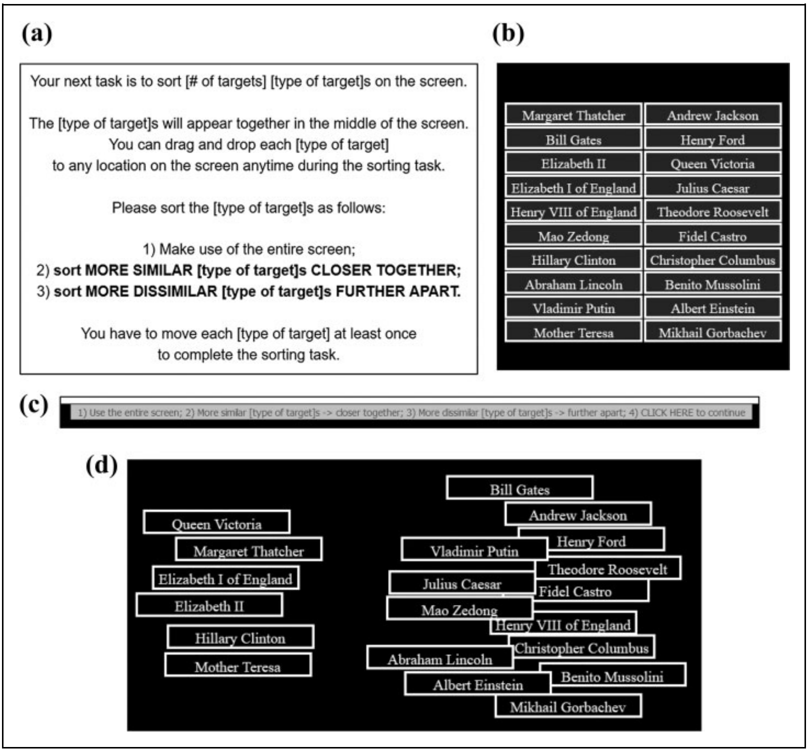

Figure 1. A. Standard instructions before spatial arrangement. B. Start of spatial arrangement method. C. Standard instructions during spatial arrangement. D. Possible solution of the spatial arrangement method task.

border and background color and the boxes' font (style, size, color, etc.). One may increase or decrease the number of columns in which the boxes appear. The boxes' initial location in their centered table is random. Instead of presenting your targets in a table (i.e., simultaneously), you can present them sequentially (on demand and in random order) in the center of the spatial arrangement slide.

Changing Spatial Arrangement Instructions

One may change the full instructions before the spatial arrangement screen (see Figure 1A) as well as shortened instructions presented throughout spatial arrangement at the bottom of the screen (see Figure 1C).

Self-(de)selecting Targets From a List

For a variety of research questions, it makes sense to measure the similarity of only those targets that participants self-select or do not self-deselect from a list prior to spatial arrangement. For example, if participants know most but not all targets in a set, self-deselecting unknown targets before spatial arrangement may be a reasonable choice, whereas if they know just a few targets, self-selecting known targets may make more sense. If one activates this feature (it is deactivated in Q-SpAM's default settings), one may require participants to self-(de)select a desired number of targets.

Self-generating Targets Based on a List

For other research questions, it makes sense to measure the similarity of only those targets that participants self-generate based on a list prior to spatial arrangement (e.g., comparing the similarity of self-generated trait inferences about the Figure 1B targets). If you activate this feature, instead of presenting for spatial arrangement only self-generated targets, you can present only the respective list prompts or self-generated targets together with their list prompts. In this case, one must increase target boxes' height because self-generated targets will appear one line below their list prompt (e.g., "Abraham Lincoln [line break] resolute").

Self-marking Targets in a List

For yet other research questions, it makes sense to measure the similarity of targets that participants self-mark in a list prior to spatial arrangement (e.g., comparing the similarity of Figure 1B targets self-marked as U.S. American and not self-marked). If one activates this feature, one may require participants to self-mark less than or equal to N, more than or equal to N, or N targets in one color or two colors. The default color is blue for the former and blue and orange for the latter; the colors can be set at will.

Changing Background of Spatial Arrangement Slide

Instead of spatial arrangement on a blank slide, one may change the background of spatial arrangement slide such that it shows coordinate axes (i.e., a horizontal and a vertical crossing in the slide's center) or circles with steadily increasing radius to the slide's center. Axes make sense if participants shall use two orthogonal dimensions (rather than, e.g., two or more clusters), whereas circles make sense if participants shall spatially arrange targets more applicable, relevant, or similar to something represented by the slide's center closer to the center (e.g., the self, some in-group, or another category such as "morality"). If one uses axes, participants can self-label the two orthogonal dimensions they used after finalizing their spatial arrangement. If one uses circles, one can change the term in the slide's center, the number of circles around it, and the steady radius increment from circle to circle.

Spatial Arrangement of Images Instead of Words

One may also present images instead of words. To do so, these images need to be in a link-accessible online directory. The respective hypertext transfer protocol link "http:// . . . " must end with "/", and the directory must contain all images. The file format of the images can be .jpg or .png, and the file name of the images has to be the same as in pools 1–10. Also, pixel width and height of the target boxes must be large enough to fit the images. Image targets are typically larger than word targets, and thus, the typical online crowdsourcing platform participant's spatial arrangement slide (i.e., something similar to 1366 by 768) fits less image compared to word targets.

Combining the Above and Other Qualtrics Features

One may combine the above features. Self-(de)selection is followed by self-generation that is followed by self-marking that is followed by spatial arrangement. Thus, participants can self-select only those Figure 1b targets they know, then self-generate a trait inference about each self-selected target, then self-mark in blue and orange all U.S. American and foreign self-selected targets, respectively, and then spatially arrange the self-selected targets self-marked as U.S. American (blue) or foreign (orange) together with the respective self-generated trait inference (e.g., "Abraham Lincoln [line break] resolute" and "Albert Einstein [line break] ingenious") along a horizontal and a vertical axis instead of on a blank slide. Further, by means of Qualtrics' built-in question types, one can instruct participants to rate, rank, and/or comment targets after spatial arrangement. For example, after self-selecting, self-generating, self-marking, and spatial arrangement combined as described above, participants may use a 0–100 slider scale ranging from "uninfluential" to "influential" to rate in random order the targets they spatially arranged just before. Alternatively, one may provide participants with multiple-choice ratings, rankings, commenting fields, and so forth.

xwindowsi	ywindowsi	Sp	label6name	label6x	label6y	Sp	label10name	label10x	label10y
904	765	...	Barack Obama	698	201	...	Bill Gates	487	259
1366	594	...	Barack Obama	213	316	...	Bill Gates	1012	165
1366	635	...	Barack Obama	1010	124	...	Bill Gates	1208	92
1309	698	...	Barack Obama	541	364	...	Bill Gates	521	129
1280	682	...	Barack Obama	78	516	...	Bill Gates	250	552
842	962	...	Barack Obama	-9999	-9999	...	leer	-9999	-9999

Figure 2. Qualtrics-spatial arrangement method data as downloaded via your Qualtrics account (each row is a participant).

Processing and Visualizing Similarity Data Collected With Q-SpAM

Next, we exemplify how to process and visualize similarity data collected with Q-SpAM using R and Excel scripts. We provide all example data, scripts, and results as well (see OSF website https://osf.io/zkfvh/). For this example, we made only two changes to the default settings of Q-SpAM. Participants would spatially arrange 50 instead of 20 targets appearing together in five instead of two columns and 10 rows in the middle of the screen. To invite participants, in project "Q-SpAM" and tab "distributions," we clicked on "Anonymous link" and pasted this link on MTurk.

We paid U.S. MTurkers U.S.$0.75 to "sort 50 influential leaders" (i.e., a random sample of 50 targets drawn from "pool 1"). Participants took $M =$ 368.61 seconds ($SD =$ 208.89) to complete this study (i.e., to spatially arrange more similar influential leaders closer together on the screen), and we recruited 99 participants (37 women, 61 men, 1 other; $M_{\text{age}} =$ 35.57, SD = 9.20) in just two hours. These numbers illustrate the efficiency of combining Q-SpAM with MTurk to collect similarity data. Before downloading these data, in project "Q-SpAM," tab "Data & analysis," tab "Add filter," question "Data_quality" (i.e., the study's last question), we selected operator "Is empty" and deleted all participants filtered in this way (i.e., those who did not complete the study). To download the remaining data, in project "Q-SpAM," tab "Data & analysis," button "Export & import," button "Export Data," "tab "XML," button "Use legacy exporter," we clicked on button "® Download," which downloaded file "Q-SpAM.xml."

Figure 2 shows that file "Q-SpAM.xml" (accessible with Microsoft Excel) saved the pixel width (first column: "xwindowsize") and height (second column: "ywindowsize") available on each participant's screen as well as the width and height coordinates of each target (e.g., columns "label6x" and "label6y" for "Barack Obama" respectively) as spatially arranged by each participant (rows 1 to *N*). Each participant's left- and uppermost coordinate was 0 and 0, respectively, and −9999 means that this target was not in the random sample to be spatially arranged by this participant. For example, participant 2 spatially arranged "Barack Obama" and "Bill Gates" to the middle-left and upper-right of the screen, respectively, whereas participant 3 spatially arranged both to the upper right of the screen. So, participant 3 (vs. 2) judged the influential leaders as more similar to each other.

Next, we describe the R script "Q SpAM.R" to compute and visualize targets' similarity as spatially arranged by each participant individually and on average (i.e., across participants). At the top of R script "Q SpAM.R," we specified "C:\\Users\\[your user name]\\Desktop\\Q-SpAM\\" as input and output folder, "Q SpAM.xml" as input file (i.e., one obviously needs to adjust these), and 62 as number of targets.

For each participant, the script saves in the specified folder the Euclidean pixel distance (i.e., dissimilarity) between each target and each other target divided by this participant's maximum possible dissimilarity (i.e., the diagonal of the participant's screen; e.g., Koch, Aleves et al. 2016; Koch et al. 2018a). The script saves one spatial arrangement visualization per participant, too. For each target pair, the script also saves dissimilarity averaged across all participants who had spatially arranged this pair. These $N \times N$ pairwise average dissimilarity indices could range from 0 to 1 (0 = *all participants spatially arranged the two targets in this pair to the same spot somewhere on their screen* and 1 = [. . .] *to opposite ends of one of their two screen diagonals*), and the script saved the indices in an $N \times N$ matrix (here: a 62×62 matrix).

Subjecting this matrix to the ALSCAL algorithm (Young, Takane, and Lewyckyj 1978; for an introduction to multidimensional scaling, see Hout, Papesh, and Goldinger 2013), the script then estimates and saves coordinates for each target in a two-, three-, four-, and five-dimensional map. We specified two- to five-dimensional map at the top of the script "dimensions<-c(2:5)"; it is possible to specify more dimensions.[3] In each map, Euclidean proximity between targets indicated that participants spatially arranged these targets closer together (i.e., had judged the targets as more similar to one another) on average. The script saves a visualization of the two- and three-dimensional map, too.

Using Q-SpAM in Social Science

Again based on the example of studying the domain of influential leaders, below we exemplify how to use Q-SpAM in social science.

Dimension- and Cluster-analyzing How People See the World

Participants must spontaneously select some dimension(s) and/or some cluster(s) to judge the similarity of some targets. For example, Barrack Obama and Bill Gates are judged as similar or dissimilar because they are seen as close or apart on some spontaneously selected dimension(s) and/or because they are seen as members of the same or different spontaneously selected cluster(s). Importantly, it is possible to reverse engineer the dimension(s) and/or cluster(s) that participants spontaneously selected to judge the similarity of some targets. Given representative participants and targets representative of some relevant domain, this research is arguably worthwhile and interesting. For example, it might contribute to understanding how people mentally organize this target domain by confirming known and exploring novel dimensions and/or clusters, their distributions, and/or their degree of relatedness. Further, it might reveal which targets are seen as low, moderate, and high on which dimensions and which targets are seen as peripheral and central in which cluster. It might confirm and explore which targets must change which dimension score and/or which cluster membership to adjust the trade-off between advantages of similarity (e.g., higher liking and cooperation) and dissimilarity (e.g., higher attention and recognition) to specific and/or all other targets, too.

To dimension- and cluster-analyze how people see the domain of influential leaders, in 2014, we paid 100 MTurkers (44 women, 56 men; $M_{\text{age}} = 33.11$, $SD = 10.18$) U.S.$1.5 to name 40 influential leaders "alive today or [. . .] passed away." We retained all 62 responses given by at least 15 percent of participants (George Washington [97 percent], [. . .], Pope Francis [15 percent]), an arguably representative sample of influential leaders. Next, we used Q-SpAM, MTurk, and the above R script to measure and process their similarity as reported above. Figure 3 (created with Microsoft Excel file "Q-SpAM.xlsx") shows their similarity parsimoniously scaled in a two-dimensional map in which higher proximity indicates higher similarity. Inspecting the horizontal, vertical, and diagonals of this map, we hypothesized that participants spontaneously selected five bipolar dimensions to judge the similarity of these influential leaders: dominant–prestigious (Cheng et al. 2013), pragmatic–idealistic, worldly–spiritual, powerful–high status, and foreign (to the United States) domestic (Zou and Cheryan 2017).

To test this, another 100 MTurkers (39 women, 61 men; $M_{\text{age}} = 32.55$, $SD = 11.27$) received U.S.$1 for using slider scales to rate the influential leaders on one of these five dimensions (e.g., "[. . .] whether they are dominant or prestigious"). We predicted the influential leaders' average ratings on

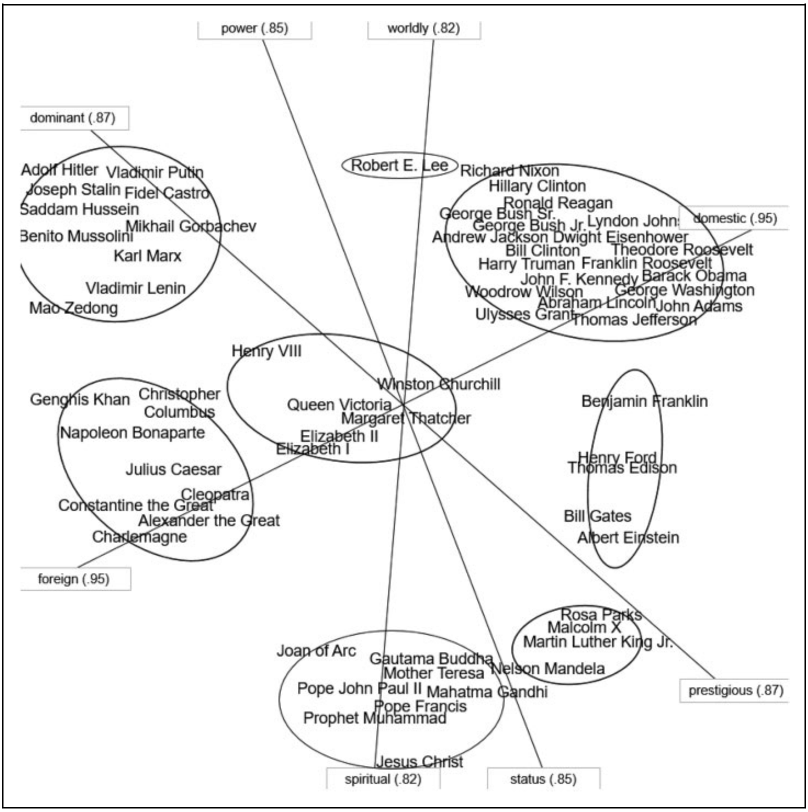

Figure 3. Using Q-SpAM.qsf, Q-SpAM.xml, Q-SpAM.R, and Q-SpAM.xlsx to dimension- and cluster-analyze how people see the domain of influential leaders.

"dominant–prestigious" from their width and height coordinates in the map. The multiple correlation of this property fitting analysis (Chang and Carroll 1969) indicated that dominant–prestigious correlated almost perfectly (r = .87) with a dimension running from the upper left to lower right of the map (see Figure 3). Thus, we concluded that participants had spontaneously selected (a synonym of) dominant–prestigious to judge the similarity of the influential leaders. Further property analyses indicated that participants had also spontaneously selected (synonyms of) worldly–spiritual (r = .82), powerful–high status (r = .85), and foreign (to the United States) domestic (r = .95) but not pragmatic–idealistic because r = .61 was substantially lower than perfect.

Table 1. Dimension-Analyzing How People See Influential Leaders.

	As Running Through the Similarity Map			At Zero Order		
Dimension	(2)	(3)	(4)	(2)	(3)	(4)
(1) Dominant–prestigious	.56	.89	.50	.52	.80	.35
(2) Worldly–spiritual		.87	.46		.72	.39
(3) Powerful–high status			−.05			−.05
(4) Foreign (to the United States)–domestic						

Table 1 shows the degree of relatedness of these spontaneously selected dimensions both at zero order and as running through (i.e., predicted from) the similarity map.

Table 2 shows the results of a subsequent cluster analysis that complemented the above insights on spontaneously perceived dimensions of influential leaders with insights on spontaneously perceived categories of influential leaders falling along these dimensions. Specifically, Table 2 displays labels and describes eight categories of influential leaders created based on the agglomeration plot and dendrogram of hierarchical clustering with between groups as linkage method and squared Euclidean as distance measure. We chose hierarchical clustering because cluster number was unknown (Yim and Ramdeen 2015).

Last, labeling a similarity map's dimensions and clusters can be left to participants, too, to further increase their say in researching how they mentally organize the respective target domain. In this way, we have shown that participants spontaneously select agency/socioeconomic success (A), conservative–progressive beliefs (B), and communion (C) to judge the similarity of societal groups (taken together: the ABC model), that participants spontaneously select agency and progressiveness (Koch, Imhoff et al. 2016) to judge the similarity of occupational groups (Imhoff et al. 2018), and that participants spontaneously select geography alongside history and agency alongside beliefs to judge the similarity of U.S. states (Koch et al. 2018a).

As these examples show (for a review, see Koch and Imhoff 2018), dimensions, clusters, their distributions, and/or their degree of relatedness vary (Abele-Brehm et al. 2019), calling for another map and other labels for every other target domain (e.g., values and goals, see Coelho et al. 2019) and, in fact, participant population (Henrich, Heine, and Norenzayan 2010; Hruschka et al. 2018) and thus a both precise and efficient similarity measure such as Q-SpAM.

Table 2. Cluster-Analyzing How People See Influential Leaders.

Cluster	Dominant–Prestigious	Worldly–Spiritual	Powerful–High Status	Foreign (to the United States)–Domestic
Presidents	In between	Worldly	Powerful	Domestic
Robert E. Lee	Dominant	Worldly	Powerful	Domestic
STEM geniuses	Prestigious	In between	High status	Domestic
Civil rights activists	Prestigious	Spiritual	High status	In between
Prophets/saints	Prestigious	Spiritual	High status	Foreign
Royals/premiers	In between	In between	In between	In between
Ancient rulers/conquerors	Dominant	In between	Powerful	Foreign
Modern rulers/revolutionists	Dominant	Worldly	Powerful	Foreign

Explaining Cognition and Behavior Based on How People See the World

Similarity profoundly influences cognition and behavior. Thus, once people's mental organization of a target domain is similarity-mapped and labeled, it is interesting to test whether its dimensions, clusters, their distribution, and/or their degree of relatedness explain downstream cognition and behavior. For example, participants spatially arranged all sorts of positive (vs. negative) people, objects, and events in a denser cluster indicating higher similarity (Alves et al. 2016; Koch, Aleves et al. 2016). Thus, we hypothesized that this valence asymmetry in targets' similarity (i.e., good is more alike than bad; for an explanation, see Alves, Koch, and Unkelbach 2017) co- or re-explains a host of cognitive and behavioral valence asymmetries.[4] And indeed, participants who spatially arranged positive (vs. negative) targets in a denser cluster later divided these into fewer categories (Koch, Aleves et al. 2016). Also due to valence asymmetry in targets' similarity, participants better recognized old versus new negative (vs. positive) targets (Alves et al. 2015), and participants more likely generalized across positive (vs. negative) personality traits when forming an impression of someone (Gräf and Unkelbach 2016, 2018).

Q-SpAM enables single-session measurement of both similarity and downstream cognition and/or behavior, which in turn enables attempts to efficiently co- or re-explain other effects in cognition and behavior in terms of valence asymmetry in targets' similarity. For example, participants were slower at (i.e., less attentive when) uttering negative (vs. positive) targets' font color (Pratto and John 1991), it took seven positive personality traits to neutralize participants' impression of someone with five negative ones (Snyder and Tormala 2017), and participants concluded performance decline based on less evidence of failure (vs. concluding performance improvement based on evidence of success; O'Brien and Klein 2017). Valence asymmetry in targets' similarity might explain these effects, too.

In other studies, politically extreme (vs. moderate) U.S. participants not only spatially arranged U.S. politicians, societal groups, and newspapers in denser clusters indicating higher similarity but also estimated that Democratic and Republican voters have less in common, meet and mingle less, and live further apart. In sum, Q-SpAM contributed to showing that "the political domain appears simpler to the politically extreme than to political moderates" (Lammers et al. 2017). Baldwin, Landau, and Swanson (2018) found another use for Q-SpAM: Participants who spatially arranged events of their biography in a denser cluster indicated higher meaning in life.

Focusing on three of Q-SpAM's nondefault features, we close by sharing ideas for using this tool in future studies. First, the self-(de)selecting feature might be useful for behavioral marketing research. Facing both liked and disliked brands and or products, valence (i.e., good vs. bad) will likely guide, if not, govern, consumers' spatial arrangement. To zoom in on, and dimension- and cluster-analyze, their mental organization of liked targets only (e.g., because they will not consider buying or recommending the disliked ones), participants should first select the ones they like and then spatially arrange only these. For example, this evasion of valence-based mental organization might reveal gaps in the market. The self-generation feature might help with studying how people mentally represent social groups versus familiar members thereof versus both, ends with versus without envisioned means, or benefits versus unpacked costs versus both. Self-generating targets might also contribute to probing lay theories of multivariate phenomena such as world order (Brandt, Sibley, and Osborne 2019) and romantic relationship. Finally, the self-marking feature might reveal insights about how people make sense of, and behave towards, friends, foes, and neutral parties tagged in blue, orange, and gray (Q-SpAM's default color).

In summary, spatial arrangement is a highly useful tool to gather similarity data for behavioral and cognitive research questions. Q-SpAM allows to collect such data online with large sets of target stimuli and large participant samples from crowdsourcing platforms. And clearly, Q-SpAM maybe used directly in the lab under more controlled circumstances as well.

Conclusion

Spatial arrangement, a task in which participants drag and drop more similar targets closer together on the screen, is a both precise and efficient similarity measure. We provided a feature-rich (self-[de]selection, self-generation, and self-marking of targets, etc.) yet user-friendly version of this spatial arrangement task built into the online survey software Qualtrics, namely Q-SpAM. We showed how to set up and customize Q-SpAM, how to measure lots of targets' similarity in little time by combining Q-SpAM with crowdsourcing platforms such as Amazon's MTurk, and how to efficiently process and visualize this similarity data with R and Excel scripts that we provide alongside with Q-SpAM. Finally, we made several suggestions how to use Q-SpAM in social science (i.e., dimension- and cluster-analyzing how people see the world and explaining downstream cognition and behavior). We believe that Q-SpAM thereby provides a tool to further illuminate the role of similarity between stimuli in cognition and behavior.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a C-SEB Junior Start-Up Grant to Alex Koch.

ORCID iD

Alex Koch https://orcid.org/0000-0002-6267-8066

Supplemental Material

Supplemental material for this article is available online.

Notes

1. We prefer the label spatial arrangement over projective mapping because (as with pairwise judgment and pile sorting) spatial arrangement describes what people do whereas projective mapping interprets what people do.
2. Or with a delay if they pass the researchers' checks for serious and attentive participation; immediate or delayed payment is up to the researchers.

3. The more dimensions in a multidimensional scaling map, the less parsimonious this map but the more variance in pairwise average dissimilarity retained in the map (a trade-off to be solved).
4. The standard explanation is negative targets' higher affective-motivational potential (i.e., bad is stronger than good; Baumeister et al. 2001).

References

Abele-Brehm, Andrea, Naomi Ellemers, Susan Fiske, Alex Koch, and Vincent Yzerbyt. 2019. "Navigating the Social World: Shared Horizontal and Vertical Evaluative Dimensions." *Under Review*.

Alves, Hans. 2018. "Sharing Rare Attitudes Attracts." *Personality and Social Psychology Bulletin* 44:1270-83.

Alves, Hans, Alex Koch, and Christian Unkelbach. 2016. "My Friends Are All Alike—The Relation between Liking and Perceived Similarity in Person Perception." *Journal of Experimental Social Psychology* 62:103-17.

Alves, Hans, Alex Koch, and Christian Unkelbach. 2017. "Why Good Is More Alike than Bad: Processing Implications." *Trends in Cognitive Sciences* 21: 69-79.

Alves, Hans, Christian Unkelbach, Juliane Burghardt, Alex Koch, Tobias Krüger, and Vaughn D. Becker. 2015. "A Density Explanation of Valence Asymmetries in Recognition Memory." *Memory & Cognition* 43:896-909.

Baldwin, Matthew, Mark J. Landau, and Trevor J. Swanson. 2018. "Metaphors Can Give Life Meaning." *Self and Identity* 17:163-93.

Bijmolt, Tammo, Michel Wedel, Rik G. M. Pieters, and Wayne DeSarbo. 1998. "Judgments of Brand Similarity." *International Journal of Research in Marketing* 15:249-68.

Brandt, Mark J. 2017. "Predicting Ideological Prejudice." *Psychological Science* 28: 713-22.

Baumeister, Roy F., Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs. 2001. "Bad is stronger than good." *Review of General Psychology* 5:323-70.

Brandt, Mark J., Chris G. Sibley, and Danny Osborne. 2019. "What Is Central to Political Belief System Networks?" *Personality and Social Psychology Bulletin* 45:1352-64.

Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, yet High-quality, Data?" *Perspectives on Psychological Science* 6:3-5.

Burger, Jerry M., Nicole Messian, Shebani Patel, Alicia del Prado, and Carmen Anderson. 2004. "What a Coincidence! The Effects of Incidental Similarity on Compliance." *Personality and Social Psychology Bulletin* 30:35-43.

Byrne, Donn. 1969. "Attitudes and Attraction." *Advances in Experimental Social Psychology* 4:35-89.

Chang, Jih-Jie and J. Douglas Carroll. 1969. *How to Use PROFIT, a Computer Program for Property Fitting by Optimizing Nonlinear or Linear Correlation*. Murray Hill, NJ: Bell Telephone Laboratories.

Cheng, Joey T., Jessica L. Tracy, Tom Foulsham, Alan Kingstone, and Joseph Henrich. 2013. "Two Ways to the Top: Evidence that Dominance and Prestige Are Distinct yet Viable Avenues to Social Rank and Influence." *Journal of Personality and Social Psychology* 104:103-25.

Coelho, Gabriel Lins de Holanda, Paul H. Hanel, Mark K. Johansen, and Gregory R. Maio. 2019. "Mapping the Structure of Human Values through Conceptual Representations." *European Journal of Personality* 33:34-51.

DeBruine, Lisa M. 2002. "Facial Resemblance Enhances Trust." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269:1307-12.

Fischer, Ilan. 2009. "Friend or Foe: Subjective Expected Relative Similarity as a Determinant of Cooperation." *Journal of Experimental Psychology: General* 138:341-50.

Gentner, Dedre and Arthur B. Markman. 1997. "Structure Mapping in Analogy and Similarity." *American Psychologist* 52:45-56.

Godwin, Hayward J., Michael C. Hout, and Tamaryn Menneer. 2014. "Visual Similarity Is Stronger than Semantic Similarity in Guiding Visual Search for Numbers." *Psychonomic Bulletin & Review* 21:689-95.

Goldstone, Robert L. 1994. "An Efficient Method for Obtaining Similarity Data." *Behavior Research Methods, Instruments, & Computers* 26:381-86.

Goldstone, Robert L., Douglas L. Medin, and Jamin Halberstadt. 1997. "Similarity in Context." *Memory & Cognition* 25:237-55.

Goldstone, Robert L. and Ji Y. Son. 2005. *Similarity*. New York: Cambridge University Press.

Gräf, Michael and Christian Unkelbach. 2016. "Halo Effects in Trait Assessment Depend on Information Valence: Why Being Honest Makes You Industrious, but Lying Does Not Make You Lazy." *Personality and Social Psychology Bulletin* 42: 290-310.

Gräf, Michael and Christian Unkelbach. 2018. "Halo Effects from Agency Behaviors and Communion Behaviors Depend on Social Context: Why Technicians Benefit More from Showing Tidiness than Nurses Do." *European Journal of Social Psychology* 48:701-17.

Griffiths, Thomas L., Mark Steyvers, and Joshua B. Tenenbaum. 2007. "Topics in Semantic Representation." *Psychological Review* 114:211-44.

Henrich, Joseph, Stephen J. Heine, and Ara Norenzayan. 2010. "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33:61-83.

Horst, Jessica S. and Michael C. Hout. 2016. "The Novel Object and Unusual Name (NOUN) Database: A Collection of Novel Images for Use in Experimental Research." *Behavior Research Methods* 48:1393-409.

Hout, Michael C. and Stephen D. Goldinger. 2015. "Target Templates: The Precision of Mental Representations Affects Attentional Guidance and Decision-making in Visual Search." *Attention, Perception, & Psychophysics* 77:128-49.

Hout, Michael C. and Stephen D. Goldinger. 2016. "SpAM Is Convenient but also Satisfying: Reply to Verheyen et al. (2016)." *Journal of Experimental Psychology: General* 145:383-87.

Hout, Michael C., Stephen D. Goldinger, and Kyle J. Brady. 2014. "MM-MDS: A Multidimensional Scaling Database with Similarity Ratings for 240 Object Categories from the Massive Memory Picture Database." *PLoS ONE* 9:e112644.

Hout, Michael C., Stephen D. Goldinger, and Ryan W. Ferguson. 2013. "The Versatility of SpAM: A Fast, Efficient, Spatial Method of Data Collection for Multidimensional Scaling." *Journal of Experimental Psychology: General* 142:256-81.

Hout, Michael C., Megan H. Papesh, and Stephen D. Goldinger. 2013. "Multidimensional Scaling." *Wiley Interdisciplinary Reviews: Cognitive Science* 4:93-103.

Hruschka, Daniel J., Douglas L. Medin, Barbara Rogoff, and Joseph Henrich. 2018. "Pressing Questions in the Study of Psychological and Behavioral Diversity." *Proceedings of the National Academy of Sciences* 115:11366-68.

Imhoff, Roland, Alex Koch, and Felicitas Flade. 2018. "(Pre)Occupations: A Data-driven Model of Jobs and Its Consequences for Categorization and Evaluation." *Journal of Experimental Social Psychology* 77:76-88.

James, William. 1890. *The Principles of Psychology*. Cambridge, MA: Harvard University Press.

Jones, Michael N. and Douglas J. K. Mewhort. 2007. "Representing Word Meaning and Order Information in a Composite Holographic Lexicon." *Psychological Review* 114:1-37.

Klohnen, Eva C. and Shanhong Luo. 2003. "Interpersonal Attraction and Personality: What is Attractive—Self Similarity, Ideal Similarity, Complementarity or Attachment Security?" *Journal of Personality and Social Psychology* 85: 709-22.

Koch, Alex, Hans Alves, Tobias Krüger, and Christian Unkelbach. 2016. "A General Valence Asymmetry in Similarity: Good Is More Alike than Bad." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42:1171-92.

Koch, Alex and Roland Imhoff. 2018. "Rethinking the Nature and Relation of Fundamental Dimensions of Meaning." Pp. 167-79 in *Agency and Communion in Social Psychology*, edited by Andrea Abele and Bogdan Wojciszke. Abingdon, UK: Routledge.

Koch, Alex, Nicolas Kervyn, Matthieu Kervyn, and Roland Imhoff. 2018. "Studying the Cognitive Map of the US States: Ideology and Prosperity Stereotypes Predict Interstate Prejudice." *Social Psychological and Personality Science* 9:530-38.

Koch, Alex, Roland Imhoff, Ron Dotsch, Christian Unkelbach, and Hans Alves. 2016. "The ABC of stereotypes about groups: Agency / socio-economic success, conservative_progressive beliefs, and communion." *Journal of Personality & Social Psychology* 110:675-709.

Kriegeskorte, Nikolaus and Marieke Mur. 2012. "Inverse MDS: Inferring Dissimilarity Structure from Multiple Item Arrangements." *Frontiers in Psychology* 3:245.

Krumhansl, Carol L. 1978. "Concerning the Applicability of Geometric Models to Similarity Data: The Interrelationship between Similarity and Spatial Density." *Psychological Review* 85:445-63.

Lammers, Joris, Alex Koch, Paul Conway, and Mark J. Brandt. 2017. "The Political Domain Appears Simpler to the Politically Extreme than to Political Moderates." *Social Psychological and Personality Science* 8:612-22.

Lê, Tâm Minh, Margot Brard, and Sébastien Lê. 2017. "Holos: A Collaborative Environment for Similarity-based Holistic Approaches." *Behavior Research Methods* 49:1597-604.

Medin, Douglas L., Robert L. Goldstone, and Dedre Gentner. 1993. "Respects for Similarity." *Psychological Review* 100:254-78.

Moelich, Erika I., Magdalena Muller, Elizabeth Joubert, Tormod Naes, and Martin Kidd. 2017. "Validation of Projective Mapping as Potential Sensory Screening Tool for Application by the Honeybush Herbal Tea Industry." *Food Research International* 99:275-86.

Montez, Priscilla, Graham Thompson, and Christopher T. Kello. 2015. "The Role of Semantic Clustering in Optimal Memory Foraging." *Cognitive Science* 39: 1925-39.

O'Brien, Ed and Nadav Klein. 2017. "The Tipping Point of Perceived Change: Asymmetric Thresholds in Diagnosing Improvement Versus Decline." *Journal of Personality and Social Psychology* 112:161-85.

Pattyn, Sven, Yves Rosseel, and Alain Van Hiel. 2013. "Finding Our Way in the Social World." *Social Psychology* 44:329-48.

Peer, Eyal, Laura Brandimarte, Sonam Samat, and Allessandro Acquisti. 2017. "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research." *Journal of Experimental Social Psychology* 70:153-63.

Perrin, Lucie, Ronan Symoneaux, Isabelle Maitre, Christian Asselin, Frédérique Jourjon, and Jérome Pagès. 2008. "Comparison of Three Sensory Methods for Use with the Napping® Procedure: Case of Ten Wines from Loire Valley." *Food Quality and Preference* 19:1-11.

Pratto, Felicia and Oliver P. John. 1991. "Automatic Vigilance: The Attention-grabbing Power of Negative Social Information." *Journal of Personality and Social Psychology* 61:380-91.

Rips, Lance J. 1989. "Similarity, Typicality, and Categorization." Pp. 21-59 in *Similarity and Analogical Reasoning*, edited by Stella Vosniadou and Andrew Ortony. Cambridge, MA: Cambridge University Press.

Risvik, Einar, Jean A. McEwan, Janet S. Colwill, Regina Rogers, and David H. Lyon. 1994. "Projective Mapping: A Tool for Sensory Analysis and Consumer Research." *Food Quality and Preference* 5:263-69.

Roediger, Henry L. and Kathleen B. McDermott. 1995. "Creating False Memories: Remembering Words Not Presented in Lists." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21:803-14.

Shaver, Phillip, Judith Schwartz, Donald Kirson, and Cary O'Connor. 1987. "Emotion Knowledge: Further Exploration of a Prototype Approach." *Journal of Personality and Social Psychology* 52:1061-86.

Shepard, Roger N. 1987. "Toward a Universal Law of Generalization for Psychological Science." *Science* 237:1317-23.

Smith, Edward E. and Steven A. Sloman. 1994. "Similarity- versus Rule-based Categorization." *Memory & Cognition* 22:377-86.

Snyder, Aaron I. and Zakary L. Tormala. 2017. "Valence Asymmetries in Attitude Ambivalence." *Journal of Personality and Social Psychology* 112:555-76.

Sofer, Carmel, Ron Dotsch, Daniel H. Wigboldus, and Alexander Todorov. 2015. "What Is Typical Is Good: The Influence of Face Typicality on Perceived Trustworthiness." *Psychological Science* 26:39-47.

Suitor, J. Jill, Shirley Keeton, and Karl Pillemer. 1995. "When Experience Counts: The Effects of Experiential and Structural Similarity on Patterns of Support and Interpersonal Stress." *Social Forces* 73:1573-88.

Tversky, Amos. 1977. "Features of Similarity." *Psychological Review* 84:327-52.

Unger, Layla, Anna V. Fisher, Rebecca Nugent, Samuel L. Ventura, and Christopher J. MacLellan. 2016. "Developmental Changes in Semantic Knowledge Organization." *Journal of Experimental Child Psychology* 146:202-22.

Unkelbach, Christian, Klaus Fiedler, Myriam Bayer, Ralf Stegmüller, and Daniel Danner. 2008. "Why Positive Information Is Processed Faster: The Density Hypothesis." *Journal of Personality and Social Psychology* 95:36-49.

Verheyen, Steven, Wouter Voorspoels, Wolf Vanpaemel, and Gert Storms. 2016. "Caveats for the Spatial Arrangement Method: Comment on Hout, Goldinger, and Ferguson (2013)." *Journal of Experimental Psychology: General* 145:376-82.

Yim, Odilia and Kylee T. Ramdeen. 2015. "Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data." *The Quantitative Methods for Psychology* 11:8-21.

Young, Forrest W., Yoshio Takane, and Rostyslaw Lewyckyj. 1978. "ALSCAL: A Nonmetric Multidimensional Scaling Program with Several Individual-differences Options." *Behavior Research Methods* 10:451-53.
Zou, Linda X. and Sapna Cheryan. 2017. "Two Axes of Subordination: A New Model of Racial Position." *Journal of Personality and Social Psychology* 112:696-717.

Author Biographies

Alex Koch is an assistant professor of behavioral science in the Booth School of Business of The University of Chicago, United States. He is interested in, and has authored articles and chapters on, social perception and cognition, content and relation of (group) stereotypes, social and economic behavior, data-driven methods, similarity (assessment), valence asymmetries in information processing, and judgments of truth.

Felix Speckmann is a PhD candidate in the Center for Social and Economic Behavior of the University of Cologne, Germany. He is interested in stereotypes and prejudice, intercultural psychology, category perception, spatial arrangement, web scraping, machine learning, and multidimensional scaling.

Christian Unkelbach is a full professor of experimental psychology in the Center for Social and Economic Behavior of the University of Cologne, Germany. He has authored articles and chapters on and edited a book on valence asymmetries, evaluative judgments, fluency/easiness effects, information ecologies, evaluative conditioning, prejudices against Muslims, and sport psychology.