Contextualizing the Cognition Crisis

Alamir Novin
University of British Columbia
1961 East Mall, Vancouver, BC, Canada
alamir.novin@ubc.ca

ABSTRACT

This workshop submission argues for a more socio-technical approach to scholarly explorations of task complexity in user search and retrieval. Doing so will lead to results that are not only more applicable outside of the lab and in the "real world," but also to results that are scientifically more 1) consistent, 2) reproducible, and 3) falsifiable. To demonstrate the importance of this approach, this submission draws on recent findings that cognitive experiments that fail to contextualize cognitive tasks lead to irreproducible results both within and outside of the lab. These same problems transfer to complex search tasks that draw from the same cognitive studies (e.g., with confirmation bias) or limit the contextualization in their own experiments. However, a distributed-cognition approach acknowledges how cognition is distributed amongst contextual factors. The author suggests that a mixed-methods approach could assist with capturing more context. Specifically, a *population-based survey experiment* may provide a straight forward method for information search and retrieval scholars to create an internal valid experimental study with some of the generalizability from survey sampling. With this in mind, the workshop suggests that instead of setting out to reproduce cognitive effects via mechanistic changes in a work task, the context of the work task should be determined first. After the context and task are determined, the mechanistic variables (e.g., interface features) for participants can be experimented with. Finally, the author draws from a long scholarly debate between various information scientists and cognitive science labs to argue that the analysis of observations can determine which cognitive effects provide the best explanatory power.

CCS Concepts

• Information systems~Query representation systems~Search interfaces

Keywords

search engine results page; cognitive bias; biases; document order; context; cognition; complex search; work tasks

Introduction

The argument that research experiments require greater social context is often perceived as an argument for a "soft-science" approach to research. However, what is often forgotten is that providing social context increases the validity of scientific principles fundamental to the hard sciences, specifically the reproducibility and falsifiability [1]. To illustrate this point, the author refers to a current scholarly debate taking place amongst scientists, often via Nature journal's magazine, that is controversially referred to as the "replicability crisis" [2]. The controversial term is directed at the finding by the Open Science Collaboration (OSC) that the vast majority of papers on cognitive biases are irreproducible by other researchers [3]. However,

Information Scientists, amongst other scholars, have objected to the methods used by the OSC. Most notably, in their paper "More on 'Estimating the Reproducibility of Psychological Science'" the authors Daniel Gilbert, Gary King, Stephen Pettigrew, and Timothy Wilson (2016) summarize their back and forth exchange with the OSC thus:

OSC: "We have provided a credible estimate of the reproducibility of psychological science."

US: "No, you haven't, because (1) you violated the basic rules of sampling when you selected studies to replicate, (2) you did unfaithful replications of many of the studies you selected, and (3) you made statistical errors."

OSC (& OTHERS): "We didn't make statistical errors."

In their latest paper, Gilbert et al., (2016) conclude: "while some colleagues have challenged our Point 3, none has challenged our Points 1 or 2, probably because it requires no special expertise to see that these points are inarguable facts" [4]. The author of this workshop will not challenge, but expand on Point 2.

This workshop submission builds on these prior studies by suggesting that scholars investigating complex search tasks can make their designs more context-based by 1) first determining the context prior to a work task. Next 2) data is collected via a mixed-methods strategy known as a *population-based survey experiment* that combines an experimental design testing the work-task with the field survey sampling methods for recruiting a population. 3) Finally, both the collected quantitative and qualitative data are analyzed in specific sequences to make inferences on the relationship between a person's context, task, and their cognition.

Background: The 'Replicability Crisis' Debate

The large analysis conducted by the OCR was the Reproducibility Project, whereby 270 researchers from academic institutions across the globe conducted a massive examination of cognitive-biases on 100 psychology studies from reputable journals. Only 39% of the studies were reproducible [3].

Reproducibility problems are not a new phenomenon [5] and they affect a variety of fields, but the "soft sciences" (e.g., social science and psychology) could be at greater risk of them because they produce the most positive results in comparison [6]. On the other hand, "hard sciences," such as physics, publish their negative results more frequently [6]. In response, a few scholars called for greater protocol in the process of cognitive research. Daniel Kahneman (2013), the respected psychologist, suggests cognitive scientists collaborate by creating a board with a protocol to oversee tests on the replicability of priming effects [7]. Bavel et al., (2016), has created a set of guidelines for scholars to provide cognitive insights, while still avoiding overtly deterministic claims [8].

CHIIR 2017 Workshop on Supporting Complex Search Tasks, Oslo, Norway.

While the Reproducibility Project findings set off a debate on the value of cognitive studies, there are two contextual points that need to be considered. The first is that the OCS also found that cognitive studies did predict the effect sizes. This point is elaborated on by Bavel et al., (2016) and Gilbert et al., (2016), so the author will not be expanding on it in this submission. Instead, the author will focus on the more salient second point: the possibility that robust contextual factors explained the low reproducibility. These two points were examined further by a large-scale analysis conducted by the Many Labs, whereby 13 studies on cognitive biases were cherry-picked (partly due to their high-likelihood for replicability) and tested by 36 academic institutions [9]. The study on 6,344 participants found that only 10 of those 13 studies were replicable. This was presented as troubling news because the 13 studies were picked due to their high chance of being reproduced [10]. Furthermore, the Many Labs scholars found that context did not play an important role in the observed results [9]. However, in response to this finding, one significant irony in the Many Labs theoretical position was raised by Bavel, Mende-Siedlecki, Brady, and Reinero (2016): "if the effects chosen for replication in these projects were predominantly effects which are *a priori* unlikely to vary by context, then it would come as no surprise that context does not predict replication success." To illustrate this point, Bavel et al., (2016) conducted an analysis of 100 replication attempts in psychology, and found "that the extent to which the research topic was likely to be contextually sensitive (varying in time, culture, or location) was associated with replication success." This finding also supports the claim by the information scientists Gilbert et al., (2016) that "*Robustness to changes in condition is important, but it is not relevant to the replicability of psychological science which is what [OCS] tried to estimate."*

Replicability of Complex Search Tasks

It is important for the sciences to be founded on the concrete principles of reproducibility and falsifiability. Because information seeking refers heavily to the cognitive sciences, it is also important for the field to recognize the challenges it faces. Furthermore, it is worth questioning whether the reproducibility of *complex* search and retrieval tasks may also be too dependent on context and cognition to make mechanistic claims that "control" for contextual factors. With complex searches, the search process is often dependent on the work task that is assigned. However, the participant's cognitive understanding of the work task also requires accounting for contextual factors.

There are a variety of definitions for both *context* and *situation* and, while scholars often use them interchangeably, they should be distinguished [11]. Although "context" is a difficult concept to define, Brenda Dervin [12] writes that there is general agreement that it accounts for "the here and the now (i.e., time and space)" (p.114). Sonnenwald distinguishes context from a *situation* by arguing that a context may contain a set of several different situations [13], p.180). The *situation* is a moment in "time-space" that works as a frame that guides the *situational factors* within it [14]. The situation can consist of several situational factors, which Hert (1997, p.21) defines as "variables which were time-space-specific". Finally, the author argues that the context, situation, and its factors are separate concepts from the *process*. The process makes up the framed interactions pieced together, which create several moments in time.

Giddens (1991) argued for a dual focus on the structures of the situation and the processes by human agents [15]. However, information seeking models that account for context often do so in one of three ways: 1) They focus on the process and *implicitly* minimize the impact of context, 2) they minimize the impact of context by *explicitly* controlling for situational factors, or 3) they acknowledge the context, but only in a positivist framework whereby all situational factors are required to be "optimal" for a process to take place. Modelling for both negative and positive outcomes accounts for the dynamics of social cognitive theories (i.e., the knowledge transfer between individuals and their surrounding society) [16]. For example, people do not only work towards positive outcomes but they also try to avoid negative outcomes [16]. In the information sciences, scholars have argued that if a problem is information intensive but requires little external input then it may be successful [17], [18], but this requires the task or problem to be designed so that external input is minimized. It also disregards how context can impact how we internalize information from our surroundings, how it affects our motivations, and how we externalize that information via communication and action. Outside of the lab, context will impact *the rationality behind* how we seek information. T.D Wilson proposed the drivers of user behaviours are found in the situation [19] and information science scholars should be researching "the PERSON performing a ROLE in an ENVIRONMENT" [20].

Distributed Cognition & Search Tasks

Distributed cognition is an expansion on T.D. Wilson's proposal because it argues cognition flows in a cyclic manner between the person's internalization of the environment and the externalization of their activities [21]. The strong influence of contextual variables has been recognized by research in various fields, such as health professionals calling for a restructuring of medical care to consider patient experience [22]–[24], education experts calling for more focus on distributed cognition [25]–[27] and businesses over the last decade who recognize that corporate culture is a greater determinant of success than strategy [28], [29]. Our behavior, rationale, and sense of self are all dependent on social components that bring value and a purpose to our everyday situations. *Socio-cultural* approaches "focus on processes of interaction of individuals with other people and with physical and technological systems" [30]. Within this framework, *distributed cognition* scholars argue that a human's cognition is distributed amongst other members in society and their non-human environment (e.g., tools, artefacts, and other objects) [21]. Therefore, these social components should be included in a researcher's analysis of human engagements, such as information seeking.

Our cognition does not reside solely within the confine of a brain, but amongst people in our surroundings and artefacts in our environment. As cognition becomes further distributed, the context of our environment becomes increasingly more important. Understanding how these environmental forces might interact requires theorizing what accounts for abstract concepts such as context and cognition. If this is the case with complex search tasks, then it should avoid distilling human information seeking into simplistic X causes Y formulas. Even when facile formulas are reproducible they can lead to poor inferences if the contextual factors influencing someone's cognition are not accounted for. For example, in the current North-American political climate it is often mistakenly thought that people on the far left-wing or far right-wing of the political spectrum do not integrate with the rest of a community because they are not aware of the facts on the other side [31]. At worst, these individuals are accused of being "irrational" or having "poor

information literacy skills." However, past and current research has demonstrated that those members on either side of the spectrum are not only *fully aware* of the information on the other side, but are often more informed than the average person [32], [33]. Individuals on the fringe of the political spectrum contextualize new pieces information retrieved to fit into the ideological frameworks of their smaller social-contexts [32]. In other words, the individual's perception of the relevancy of information is dependent on the individual's relationship with smaller social groups (i.e., where a member has strong-ties) more so than their greater community (i.e., where a member has weak-ties). Thus, while a scientific observer might be technically correct that the far left or far right seeks information that is regarded as "false" by the greater community of which both the observed individual and observer are members, in certain cases it may be more accurate to say that the individual is seeking information that fits within the distributed cognition of smaller groups within the community first. However, the scientific observer is often only aware of contextual factors of a person's greater community and less aware about their smaller, more private, communities (e.g., marginalized far-right communities). Individuals often do not disclose their associations with groups that are perceived as *less rational* in a community. For example, researchers found that on online dating websites "both men and women of all political persuasions act as if they prefer same-race relationships even when they claim not to" [34]. In this scenario, it is clear that online daters are hiding their less socially-acceptable preferences for people from the same-race. However, if the scientist observes a person seeking information that is not relevant to the greater community but *unaware* of how the information is relevant to a smaller group, they will categorize this behavior as poor "information literacy." Maintaining strong ties with your social group could be considered a cognitively "rational" decision – even if it seems irrational to the larger community.

How to Test for Contextual Influences on Rational and Cognition

First, the approach researchers use to choose which contextual and cognitive variables to control for *is* a complex search task for the researchers themselves. My argument for the need to first understand the context of the work task prior to designing the task still applies. Before choosing the variables to research, researchers should ask what is the context of the research's purpose for the field? For example, is the research to strengthen reproducibility of the field or is it to seek out new theories? As I will explain in my second point: how much a science study should control context to increase reproducibility *depends* on the context of the science field.

Second, John W. Creswell has written a fairly clear explanation about how conducting a qualitative analysis before a quantitative analysis can lead to different inferences from a study that performs mixes the analysis in the opposite sequence [35]. To briefly summarize Creswell, the main question a researcher should ask is whether the researcher intends to conduct an *exploratory* or *explanatory* investigation. to measure the usefulness of information a quantitative and qualitative approach should be applied [36]. A mixed methods design may be best suited to draw inferences to explain the observed data. In addition, mixed methods can increase the responsiveness of participants in an otherwise complex research activity that is asking for a significant amount of effort [35]. More importantly, mixed methods provides an opportunity for participants to *explain* to the observer the rational behind their choices. This may seem clear, but rarely has the author observed scholars explain the reasons behind the chosen *sequence* for data collection and analysis and how the qualitative and quantitative data is "mixed" *(i.e., Creswell's explanatory or exploratory).*

Third, mixing the methodology of surveys and experiments can increase the likelihood of capturing contextual factors. Research in information search and retrieval requires accounting for the interactions users have with systems [37], the usefulness of the information [38], its influences [36], and the outcome of the retrieval [39]. Pia Borlund argues researching these factors requires tailoring the work task to the information environment and participant: "if the evaluation takes place by involvement of university students then the simulated work task situation should be to describe a situation they can relate to, and report on how the situation was simulated" [39]. To create such a situation, the task should be piloted and the final report should explain how the situation was simulated [39]. In addition, studies can simulate the context in information search and retrieval further by conducting a *population-based survey experiment:* "a population-based experiment uses survey sampling methods to produce a collection of experimental subjects that is representative of the target population of interest for a particular theory, whether that population is a country, a state, an ethnic group, or some other subgroup. The population represented by the sample should be representative of the population to which the researcher intends to extend his or her findings" [40]. This method requires participants to be randomly assigned to a situation and can take place outside of a lab and within the population itself, like field studies [40]. Diana C. Mutz argues the main advantage of choosing this method is that "theories can be tested on samples that are representative of the populations to which they are said to apply" [40]. By doing so, both the internal and external validity of an experiment can increase [40]. For example, in Borlund's example on university students, this can be accomplished by choosing the common area of a university library to study and recruiting students within their own familiar environment for an information seeking experiment. The ability of population-based survey experiments to be carried outside of the lab and directly in the field means that experiments can capture more contextual factors "in the wild."

Conclusion

In summary, to account for both the cognitive and contextual factors, the author proposes a more distributed-cognition approach. In light of the "cognitive crisis", instead of setting out to reproduce cognitive effects on an interface, the author suggests addressing this issue from another angle: the context of the work task should first be determined and then the work task. After these two realms are considered, the mechanistic variables (e.g., such as interface features) for participants can be experimented with. Cognition is dependent on context – an area scholars in distributed cognition explores more fully. Critics in favor of reproducibility projects may argue that this approach abstracts the concept of "cognition," but that criticism begs the question of whether a concrete framework of cognition-without-context even exists.

References

[1] L. Chi Vo, E. Mounoud, and J. Rose, "Dealing with the opposition of rigor and relevance from Dewey's pragmatist perspective," *M@n@gement*, vol. 15, no. 4, p. 368, 2012.

[2] E. Yong, "Psychologists strike a blow for reproducibility," *Nat. News*, Nov. 2013.

[3] Open Science Collaboration, "Estimating the reproducibility of psychological science," *Science*, vol. 349, no. 6251, pp. aac4716-aac4716, Aug. 2015.

[4] D. T. Gilbert, G. King, S. Pettigrew, and T. D. Wilson, "More on 'Estimating the Reproducibility of Psychological Science,'", 2016.

[5] T. D. Sterling, "Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance--Or Vice Versa," *J. Am. Stat. Assoc.*, vol. 54, no. 285, p. 30, Mar. 1959.

[6] D. Fanelli, "'Positive' Results Increase Down the Hierarchy of the Sciences," *PLoS ONE*, vol. 5, no. 4, p. e10068, Apr. 2010.

[7] E. Yong, "Nobel laureate challenges psychologists to clean up their act," *Nature*, Oct. 2012.

[8] J. J. Van Bavel, P. Mende-Siedlecki, W. J. Brady, and D. A. Reinero, "Contextual sensitivity in scientific reproducibility," *Proc. Natl. Acad. Sci.*, vol. 113, no. 23, pp. 6454–6459, Jun. 2016.

[9] R. A. Klein *et al.*, "Investigating variation in replicability," *Soc. Psychol.*, 2014.

[10] E. Yong, "Psychologists strike a blow for reproducibility," *Nature*, Nov. 2013.

[11] C. Cool, "The concept of situation in information science," *Annu. Rev. Inf. Sci. Technol.*, vol. 35, pp. 5–42, 2001.

[12] B. Dervin, "Given a context by any other name: Methodological tools for taming the unruly beast," *Inf. Seek. Context*, vol. 13, p. 38, 1997.

[13] D. H. Sonnenwald, "Evolving perspectives of human information behavior: Contexts, situations, social networks and information horizons," presented at the Exploring the contexts of information behavior: Proceedings of the Second International Conference in Information Needs, 1999.

[14] C. A. Hert, *Understanding Information Retrieval Interactions: Theoretical and Practical Implications*. Greenwood Publishing Group, 1997.

[15] A. Giddens, *Modernity and self-identity: Self and society in the late modern age*. Stanford University Press, 1991.

[16] C. Abraham and S. Michie, "A taxonomy of behavior change techniques used in interventions.," *Health Psychol.*, vol. 27, no. 3, p. 379, 2008.

[17] E. Kuiper, M. Volman, J. Coiro, M. Knobel, C. Lankshear, and D. Leu, "The Web as a source of information for students in K–12 education," *Handb. Res. New Literacies*, vol. 5, pp. 241–266, 2008.

[18] A. W. Lazonder, "Do two heads search better than one? Effects of student collaboration on web search behaviour and search outcomes," *Br. J. Educ. Technol.*, vol. 36, no. 3, pp. 465–475, 2005.

[19] T. D. Wilson, "Activity theory and information seeking," *Annu. Rev. Inf. Sci. Technol.*, vol. 42, no. 1, pp. 119–161, 2008.

[20] T. D. Wilson, "Information behaviour: An interdisciplinary perspective," *Inf. Process. Manag.*, vol. 33, no. 4, pp. 551–572, Jul. 1997.

[21] G. Salomon, Ed., *Distributed cognitions: psychological and educational considerations*, 1st pbk. ed. Cambridge [Cambridgeshire] ; New York: Cambridge University Press, 1997.

[22] A. Donabedian, "Evaluating the quality of medical care," *Milbank Mem. Fund Q.*, pp. 166–206, 1966.

[23] G. Gardner, A. Gardner, and J. O'Connell, "Using the Donabedian framework to examine the quality and safety of nursing service innovation," *J. Clin. Nurs.*, vol. 23, no. 1–2, pp. 145–155, Jan. 2014.

[24] S. W. Glickman, K. A. Baggett, C. G. Krubert, E. D. Peterson, and K. A. Schulman, "Promoting quality: the health-care organization from a management perspective," *Int. J. Qual. Health Care*, vol. 19, no. 6, pp. 341–348, Sep. 2007.

[25] Brown, L. Ann et al., "Distributed Expertise in the Classroom," in *Distributed cognitions: psychological and educational considerations*, G. Salomon, Ed. Cambridge [England]; New York, NY: Cambridge University Press, 1993, pp. 188–228.

[26] E. M. Meyers, "Mediating group search: Lessons from a middle school study," *Teach. Libr.*, vol. 38, no. 2, p. 24, 2010.

[27] E. M. Meyers, K. E. Fisher, and E. Marcoux, "Making Sense of an Information World: The Everyday-Life Information Behavior of Preteens," *The Library*, vol. 79, no. 3, 2009.

[28] J. H. Dyer and K. Nobeoka, "Creating and managing a high-performance knowledge-sharing network: the Toyota case," *Strateg. Manag. J.*, vol. 21, no. 3, pp. 345–367, Mar. 2000.

[29] T. Kell and G. Carrott, "Culture matters most," 2005.

[30] A. Collins, J. Greeno, L. Resnick, B. Berliner, and R. Calfee, "Cognition and learning," *B Berl. R Calfee Handb. Educ. Psychol. N. Y. Simon Shuster MacMillan*, 1992.

[31] H. Landemore, "On Minimal Deliberation, Partisan Activism, and Teaching People How to Disagree," *Crit. Rev.*, vol. 25, no. 2, pp. 210–225, Jun. 2013.

[32] D. M. Kahan, A. R. Landrum, K. Carpenter, L. Helft, and K. H. Jamieson, "Science Curiosity and Political Information Processing," *Adv. Polit. Psychol. Forthcom.*, 2016.

[33] D. C. Mutz, "Reflections on *Hearing the Other Side* , in Theory and in Practice," *Crit. Rev.*, vol. 25, no. 2, pp. 260–276, Jun. 2013.

[34] A. Anderson, S. Goel, G. Huber, N. Malhotra, and D. Watts, "Political Ideology and Racial Preferences in Online Dating," *Sociol. Sci.*, pp. 28–40, 2014.

[35] K. B. Rasmussen, "General Approaches to Data Quality and Internet-generated Data," in *The Sage Handbook of online Research Methods*, Fielding, Lee, and Blank, Eds. Sage, 2008.

[36] D. Case, "Sixty years of measuring the use of information and its sources: from consultation to application," *Libr. Digit. Age LIDA Proc.*, vol. 13, 2014.

[37] D. Kelly, "Methods for evaluating interactive information retrieval systems with users," *Found. Trends Inf. Retr.*, vol. 3, no. 1—2, pp. 1–224, 2009.

[38] N. J. Belkin, M. Cole, and J. Liu, "A model for evaluation of interactive information retrieval," presented at the Proceedings of the SIGIR 2009 Workshop on the Future of IR Evaluation, 2009, pp. 7–8.

[39] P. Borlund, "Interactive Information Retrieval: An Introduction," *J. Inf. Sci. Theory Pract.*, vol. 1, no. 3, pp. 12–32, Sep. 2013.

[40] D. C. Mutz, *Population-based survey experiments*. Princeton University Press, 2011.