To Train or Not to Train? How Training Affects the Diversity of Crowdsourced Data

Completed Research Paper

Shawn Ogunseye
Bentley University
Massachusetts, USA
sogunseye@bentley.edu

Jeffrey Parsons
Memorial University of Newfoundland
Newfoundland, Canada
jeffreyp@mun.ca

Roman Lukyanenko
HEC Montréal
Quebec, Canada
roman.lukyanenko@hec.ca

Abstract

Organizations and individuals who use crowdsourcing to collect data prefer knowledgeable contributors. They train recruited contributors, expecting them to provide better quality data than untrained contributors. However, selective attention theory suggests that, as people learn the characteristics of a thing, they focus on only those characteristics needed to identify the thing, ignoring others. In observational crowdsourcing, selective attention might reduce data diversity, limiting opportunities to repurpose and make discoveries from the data. We examine how training affects the diversity of data in a citizen science experiment. Contributors, divided into explicitly and implicitly trained groups and an untrained (control) group, reported artificial insect sightings in a simulated crowdsourcing task. We found that trained contributors reported less diverse data than untrained contributors, and explicit (rule-based) training resulted in less diverse data than implicit (exemplar-based) training. We conclude by discussing implications for designing observational crowdsourcing systems to promote data repurposability.

Keywords: Data repurposability, information diversity, data-driven insight, crowdsourcing

Introduction

Crowdsourcing is a popular way of outsourcing tasks normally done by an organization or by professionals to an undefined, frequently online, group with varying levels of motivation and skill. One popular form of crowdsourcing is ***observational crowdsourcing*** – engaging people to provide data based on experiences and observations, such as the migratory patterns of wildlife (Lukyanenko and Parsons 2019) or the surveillance of potential disease outbreaks (Palmer et al. 2017).

A key challenge in crowdsourcing is data quality. One popular strategy to ensure data quality in observational crowdsourcing is to recruit contributors who have the knowledge necessary to perform the data collection task (Surowiecki, 2005; Wiggins & He, 2016; Wiggins et al., 2011). When this is not possible (e.g., knowledgeable contributors are scarce or cannot be readily identified or targeted), an alternative strategy is to train potential contributors to attain the desired level of proficiency in the data reporting task (Yang et al. 2018). Training teaches contributors to provide data accurate and complete enough to be useful for the immediate purpose for which a data collection task was designed.

However, crowdsourced data are increasingly used in ways not envisioned when the data were collected. For example, Yelp review data, which is intended to guide shoppers and merchants on the weaknesses and strengths of services provided by businesses, has been used to identify restaurants with a high risk of health code violation and outbreaks of foodborne diseases (Harrison et al. 2014; Nsoesie et al. 2014; Schomberg et al. 2016). Similarly, mobile check-in data from Foursquare, an app used to share location information with friends and family, and Yelp reviews have been used to accurately predict business failures (Wang et al. 2015). Indeed, the rise of data science is largely predicated on *repurposing data* - using specific datasets in ways that go beyond the intended uses anticipated when the data collection process was designed. As repurposing is inherently unpredictable, this raises an important question about how to support repurposability when designing a data collection platform and process.

One important characteristic that makes data repurposable is its ***diversity*** – the extent to which records in a data set contain information about different features of the observed phenomena (Ogunseye & Parsons 2018). A data collection process that generates the same features for each new observation that is reported has low diversity. This is ideal when the features are those needed for the task for which the data collection process was designed. In contrast, a process that generates different features for each observation has high diversity. Such data might contain data useful for purposes not unanticipated at the time the data were collected. Diverse data, reflecting the different perspectives of contributors, can better meet the requirements of different users and different uses, even lead to discoveries (Ghasemaghaei & Calic, 2019; Parsons & Wand, 2014; Woodall, 2017).

This research focuses on the effects of training on the diversity of data collected in observational crowdsourcing settings. Training helps contributors reliably provide data that meet anticipated uses (e.g., identifying a species of organism in a biology-focused citizen science application). However, we do not know how training affects the diversity of data collected in observational crowdsourcing. We conducted a lab experiment asking people to report sightings of artificial entities to test the effect of explicit (rule-based) training and implicit (exemplar-based) training on data diversity. We found that the diversity of data collected depends on whether and how we train crowd members, with untrained contributors providing the most diverse data and explicitly trained contributors providing the least diverse data.

Next, we describe the theoretical foundations for our study and propose several hypotheses about the effects of training on diversity. We then describe our experiment, present the results, and discuss the implications of our findings.

Theoretical Foundations

Relevant Ontological Concepts

To better understand the concept of data diversity, we consider how the real world is represented using data. Bunge's ontology helps us understand the structure of the real world (Shanks et al. 2008; Wand and Weber 1990) and provides a useful framework with which to explore how contributed data represents the real world. Bunge's ontology posits that the world is made up of things – unique and substantial individuals (Bunge, 1977. Things can be composed of other things (parts). Humans perceive the attributes of things and identify things using these attributes. We identify the *Aedes aegypti* mosquito*,* which transmits diseases such as dengue hemorrhagic fever, yellow fever, and Zika virus, using its attributes: "*A. aegypti* is easily recognized by the contrasting black and white rings on its legs and the lyre-shaped pattern of silver markings on the upper surface of the thorax" (Rozendaal et al. 1997, p. 16). The attributes of a thing can either be ***intrinsic***, i.e., solely depending on the thing, or ***mutual***, i.e., dependent upon more than one thing (Bunge 1977; Parsons and Wand 2000; Wand and Weber 1990). The black and white rings on the legs of the *Aedes aegypti* are an *intrinsic attribute*. In contrast, where the mosquito *dwells* (e.g., in a water storage tank*)* is a mutual attribute depending on two things: the mosquito and the tank. Mutual attributes inform us about the behavior, actions, and relationship between things (Kiwelekar and Joshi 2010; Rosemann and Green 2002; Wand et al. 1999).

Hypotheses Development

How contributors learn

During training, data contributors learn what attributes are essential for identifying a thing, that is, the diagnostic attributes of a thing. Contributors are expected to prioritize and direct their attention to these attributes when classifying the thing (Buschman and Miller 2007, p. 1860). If a contributor is not trained and has no prior knowledge of a thing, attention is driven by the salient attributes of the thing (Buschman & Miller, 2007), and by the cognitive effort required to search for attributes. Salient attributes are those that are prominent in a contributor's visual space, such as the color, size, and shape of the parts of a thing. Salient attributes attract untrained observers' attention more than they do trained observers (Theeuwes 2010; Buschman and Miller 2007; Katsuki and Constantinidis 2014). For example, infants (six to eight months) and young children who lack prior knowledge tend to pay attention to more of a thing's attributes than adults. Adults tend to pay attention to a few specific attributes of a thing needed to classify or identify the thing (Best et al. 2013; Gelman and Markman 1986; Kloos and Sloutsky 2008).

Training leads to two common forms of knowledge: implicit and explicit knowledge (Berry and Dienes 1993; Sun et al. 2005; Vargios 2007). ***Implicit knowledge*** is experiential knowledge acquired by observing the structure or attributes of a thing, while ***explicit knowledge*** is acquired through instruction or rules and can be more readily verbalized than implicit knowledge (Leonardi and Bailey 2008)[1]. Acquiring implicit knowledge (implicit learning) occurs naturally and can be inadvertent as people gain experience in a task (Reber 1993; Taylor 2007). Teaching contributors by exposing them to a class of things allows contributors to infer the diagnostic attributes of the class. Implicitly trained contributors infer diagnostic attributes by observing the similarities in attributes between members of the class (Rosch 1973). When implicitly trained, contributors acquire their knowledge using bottom-up attentional allocation by attending to salient attributes of the thing they observe and thus learning the different possible diagnostic attributes of the thing (Ogunseye et al. 2017).

Explicit knowledge is acquired through rule-based training (explicit training) – teaching contributors the rules needed to identify or categorize a thing. It is acquired consciously. By training contributors explicitly, the trainer provides the contributor with the rules that govern the identification of a thing and membership of a class (Jensen et al. 2017). These rules consist of a set of attributes, which are usually a proper subset of all the thing's attributes.

How contributors use knowledge

When observing a thing, many attributes of both the thing and its surroundings compete for an observer's attention. Having learned the diagnostic attributes of the thing, trained contributors will focus on those attributes that are needed to classify the thing, i.e., identify and make inferences about the thing (Bjorklund and Harnishfeger 1990). This helps manage limited cognitive resources, but in doing so, trained contributors will ignore attributes of the primary thing and any other things in their perceptual field that are irrelevant to classifying the thing (Hoffman and Rehder 2010; Prat-Ortega and de la Rocha 2018). This phenomenon is called ***selective attention***: "the differential processing of simultaneous sources of information" (Johnston and Dark 1986, p. 44). The sources of information competing for an observer's attention can be auditory, visual, or memorial (Plude et al. 1994). What aspects of these competing sources an observer pays attention to determine what they remember and report as data (Neill et al. 1995).

When attributes that have been learned are used to guide attention, then attention is directed from the top-down or is knowledge-driven (Katsuki and Constantinidis 2014). Top-down attention allocation is based on "volitional shifts of attention…derived from knowledge about the current task (e.g., finding your lost keys)" (Buschman and Miller 2007, p. 1860). However, when an observer has not previously committed the attributes of a thing to memory or has a first-time encounter with the thing, the attributes of the thing solely direct their attention, and their attention is allocated from the bottom-up or is stimulus-driven. In bottom-up attentional allocation, "target stimuli 'pop out' if they differ sufficiently from their background in terms

[1] We exclude a third type of knowledge – tacit knowledge, i.e. just unexplainably knowing how to do something, intuition – because it is difficult or impossible to teach tacit knowledge to other people (Leonardi and Bailey 2008).

of features such as color or orientation" (Katsuki & Constantinidis, 2014, p 509). Bottom-up attention is "automatic" and driven by "properties inherent in stimuli … (e.g., a flashing fire alarm)".

After training, an observer becomes more inclined to selectively attending to information as a coping mechanism to deal with the deluge of information sources competing for their attention (Richards & Turner, 2001). For example, infants cannot selectively attend to the attributes of things. Instead, their attention is driven by the salience of attributes. Adult humans decide which attributes of things to attend to based on their prior experience with similar stimuli, and they continue to value the usefulness of those attributes the more they are exposed to similar stimuli (Gazzaley and Nobre 2012). This infants-adults distinction thus exemplifies how humans allocate our attention to a thing when we lack prior knowledge and when we have prior knowledge about the thing (Keil 1989; Kloos and Sloutsky 2008). Trained contributors, like adult humans, will selectively attend to specific attributes of things, reporting only these attributes, whereas untrained contributors are less inclined to attend selectively to few attributes.

Training works because it controls what we pay attention to about an observed thing and what we report about it. When trained explicitly, an observer's view of a thing is bounded by the trainer's view of the thing presented to them, but when trained implicitly, the observer's view is limited to the available exemplars from which they have learned. Since what we selectively attend to is informed by how we have acquired knowledge, we, therefore, hypothesize about how selectively attending to attributes, based on implicit or explicit training, will affect contributors' capacity to report diverse data, variations in the attributes of things, and quality data.

Reporting Diverse Data

Diverse data is data that represents the different attributes of a thing, and the different knowledge of data contributors about the thing. It has clear benefits for data repurposability and novel discoveries or insight. For example, eighty percent of surveyed organizations that accessed diverse data from different internal sources and other businesses gained more insights than organizations that used narrowly focused data from functional silos (Ransbotham and Kiron 2017). This ability for diverse data to be repurposed makes it more valuable for data-driven decision making than homogeneous data (Günther et al. 2017).

Data diversity is a consequence of contributors being able to report data from their individual perspectives. Diversity is the total number of unique attributes reported about a thing (Ogunseye and Parsons, 2018). Data reported for an observation can be described in terms of attributes. Therefore, if A and B are observations that contributors provide about a particular observed thing, A is *more diverse* than B if A has more unique attributes about the phenomenon than B.

Training determines the overall diversity of collected data in observational crowdsourcing. It prepares contributors to focus on some things (or some attributes of things) and ignore other things (or other attributes of these things) in a visual space. For example, training primes contributors to pay more attention to a primary thing, among other perceivable secondary things. When things for which contributors have received training are interacting with secondary things in their environment, trained contributors are expected to report fewer attributes of such secondary things or their interactions than untrained contributors. If these associations with secondary things become important for another use of the data, that knowledge would have been lost. Consider an observational crowdsourcing task focused on reporting the sighting of mosquitoes. Some mosquitoes are known to live in containers (Ritchie et al., 2003; Sprenger, 1987). In a reporting task about mosquitoes, trained contributors might focus on identifying the thing, and not mention that it was found in a container. Given such data, repurposing the data for another use case where the data user needs to understand which mosquitos are likely Zika virus-transmitting types[2] becomes infeasible without redoing the data collection task.

The type of training contributors receive will also have an impact on the diversity of data they provide. Explicitly trained contributors would be less likely than implicitly trained contributors to report attributes of a primary thing and attributes of secondary things outside the finite attributes of the primary thing they have been taught to prioritize. We predict that implicitly trained contributors who have been repeatedly exposed to a primary thing during training will also pay more attention to the primary thing than to other

[2] The Aedes aegypti mosquito, which transmits the Zika virus, is a "container-breeding mosquito" ("Zika, Mosquitoes, and Standing Water | | Blogs | CDC" 2016)

things. We expect that implicitly trained contributors will attend to the intrinsic attributes of the primary thing. However, being more flexible in allocating attention than explicitly trained contributors, implicitly trained contributors will also pay more attention to the salient attributes of secondary things.

Untrained contributors are not guided by prior knowledge. What they report is from bottom-up attentional allocation. They do not know if interactions and secondary things are unimportant for the crowdsourcing task. The salience of the attributes they observe directs their attention. Consequently, we expect untrained contributors to report more attributes than explicitly trained contributors who focus on a subset of the available attributes of a primary thing. Also, unlike implicitly trained contributors who will prioritize the attributes they have learned, untrained contributors will distribute their attention to all salient attributes of both primary and secondary things in their visual field. Therefore, we predict that untrained contributors will report more attributes in aggregate about the things in their visual field than explicitly or implicitly trained contributors.

H1: Information diversity. *Untrained contributors will report more diverse data than will implicitly trained contributors, who, in turn, will report more diverse data than will explicitly trained contributors*.

Reporting Attribute Variations

The potential existence of instances whose attributes (e.g., their color, number, shape or size) deviate from the attributes to which trained contributors were exposed in training is most prevalent when crowdsourcing for information about living organisms or phenomena where human knowledge is limited. Not all available instances of an entity can be shown to potential contributors during training. Some of the attributes that contributors encounter in a crowdsourcing task may be new to only the contributor. However, in some cases, contributors may encounter attributes that are new both to themselves and to the data users. For instance, while classifying galaxies from images, a data contributor to the GalaxyZoo project – Hanny van Arkel – helped identify a "brand new type of astronomical object previously unknown to scientists" because she flagged some of its attributes as atypical; "it appeared as a blue squiggle" ("Hanny's Voorwerp – History of a Mystery" 2013). How training affects a contributor's ability to report attributes they encounter that are new to them and may lead to the discovery of new things or new states of old things is pertinent in observational crowdsourcing where discoveries are possible and welcomed (Ogunseye and Parsons 2016).

Training can limit contributors' reporting of variations to intrinsic attributes of primary things for which they learned not to attend to (Hoffman and Rehder 2010). Learning to ignore some attributes (a consequence of selective attention) leads to inattentional blindness - when a contributor fails to see some visible attributes of an entity in their visual field because they are attending to other attributes of the entity (Simons, 2000). On the one hand, we expect contributors attending to some attributes to be blind to variations in other attributes when the varying attributes have no bearing on the task (Simons and Ambinder 2005; Simons and Rensink 2005). On the other hand, we expect contributors trained to selectively attend to diagnostic attributes to report more variations affecting these diagnostic attributes when they occur.

Knowledgeable contributors will be more resistant to learning something new (Plebanek & Sloutsky, 2017), impeding their ability to report attributes they have not observed before. This will, however, differ by the type of knowledge a contributor has. Even though things can exhibit mutual and intrinsic attributes that are new to a contributor, we restrict our notion of attribute variations to significant additive differences in the intrinsic attributes of a thing *compared to the instances introduced to the contributor during training*. Explicitly trained contributors are expected to report fewer occurrences of intrinsic attribute variations than implicitly trained contributors. This is because implicitly trained contributors have attended to more intrinsic attributes while developing their classification rules (Vargios 2007), even those not required to classify the thing, and should be more sensitive to manifestations of attributes they have not seen before.

Every attribute of an unknown thing is new to an untrained contributor. Untrained contributors are therefore expected to report the attributes they observe driven by attribute salience. However, because their attention is more dispersed across their visual field, their inattentional blindness would be due to allocating their attention to competing salient attributes, and not due to a conscious focus on only known attributes. If the attribute is salient, the chance that an untrained contributor will report it is thus high. A trained contributor may still ignore an attribute variation, even if it is salient. Again, consider our mosquito example: assume the key diagnostic attribute of the *Aedes aegypti* is black and white rings on its legs. A trained contributor is more likely to report an instance of a mosquito with pale pink (or light gray) and black rings on its legs, mentioning the new color than to report the presence of an extra wing if they have learned that the number of wings is not diagnostic, even if the wings have salient attributes. An untrained contributor would more likely report the increase in wings than report the variation in ring color, assuming the latter attributes are not salient but are diagnostic.

Untrained contributors are expected to report fewer intrinsic attribute variations than implicitly trained contributors. However, we expect untrained contributors to report more attribute variations than explicitly trained contributors, especially if the new attributes do not affect the classification task.

H2: Attribute variations. *Explicitly trained contributors will report fewer attribute variations of observed instances than will untrained contributors who, in turn, will report fewer attribute variations than implicitly trained contributors*

Reporting Quality Data

Contributed data should represent the state of an observed world at a given time and should help information users reproduce that state whenever necessary. In the literature and practice, data quality is judged by the extent to which the data fit intended use (Sadiq and Indulska 2017; Wang and Strong 1996). Data users worry about possible deficiencies in their data (Forbes, 2018; Weigelhofer & Pölz, 2016). The main deficiencies data users concern themselves about are inaccuracy, i.e., when data represents a "real-world state different from the one that should have been represented" (Wand and Wang 1996, p. 93), and to a lesser extent, incompleteness – when some "lawful states of the real-world system [are not] represented by the [data]" (Wand and Wang 1996, p. 91; Wang and Strong 1996).

When data contain all the attributes of a thing that are relevant to a specific context of use or view of the thing, the data is said to be complete (Nelson et al. 2005; Wixom and Todd 2005). If the use case or view of the thing changes (i.e., data is repurposed), the data may no longer be complete, prior classifications may no longer be accurate. For example, based on a view of what constitutes a planet, Pluto was previously classified as one of the nine planets in our solar system. When the planet class was redefined, the previously sufficient attributes of Pluto became incomplete. It would now be inaccurate to classify Pluto as a planet in our solar system. What constitutes accurate and complete information (quality data) can change when the use of the data changes (Ogunseye and Parsons 2018).

Unlike the concept of data diversity, information quality is tied to an imminent use of the data. Diversity, on the other hand, accommodates repurposability, discoveries, and unanticipated uses. To illustrate the difference between data quality and diversity, consider a case where a contributor C1 reports Attributes 1 and 2 about a thing in the real world (RW), and if these attributes are considered sufficient for the task at hand by information user U1, the information is complete according to the traditional definition of completeness and assuming accurate operationalization of the complete attributes, the data is considered high quality (see Figure 1a). However, if Attribute 4 of the thing becomes relevant in the future, a new information-gathering process would have to be initiated, or novel insights may be forfeited (Bonter and Cooper 2012). Figure 1b shows that if information diversity is encouraged, such that contributors C1 and C2 provide different perspectives to the information source, consumers U1 and U2 can derive multiple views from the data.

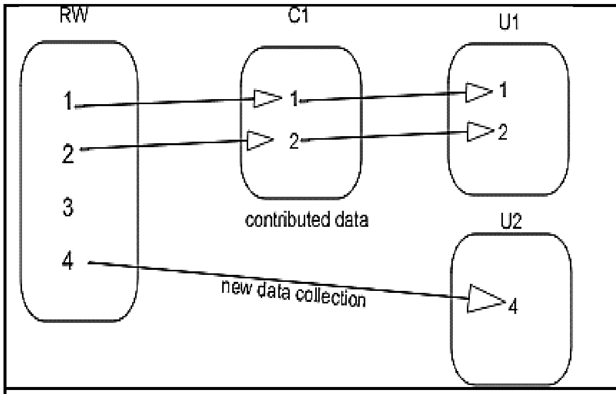


Figure 1a: Data Quality and Repurposability

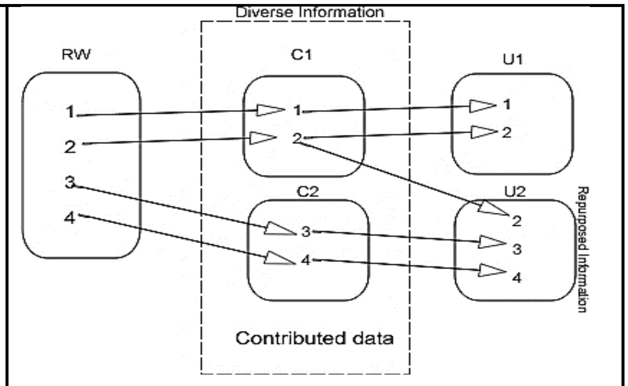


Figure 1b: Data Diversity and Repurposability

We discuss how training impacts data quality in terms of data quality deficiencies: inaccuracy and incompleteness. Our view of incompleteness focuses on the completeness of attributes rather than the completeness of things. From this view, we agree with Wand and Wang (1996) that identification made from incomplete attributes can lead to inaccuracy. The complete attributes required to identify a thing are the diagnostic attributes of the thing. Accuracy, in the context of observing and reporting phenomena, depends on the ability of contributors to map the complete attributes (diagnostic attributes) in data to the correct real-world thing.

The diagnostic attributes of things can include both their mutual attributes and their intrinsic attributes. Even though mutual attributes may provide additional criteria for identifying a thing, we assume the mutual attributes of a thing are not mandatory attributes but complementary attributes, useful for validation of accuracy. Therefore, only the intrinsic attributes of the thing are diagnostic. For instance, *Aedes aegypti* mosquitoes are expected to breed in containers and feed in the morning and at night (Rozendaal et al. 1997). However, violating these mutual attribute requirements should not invalidate the mosquito's class.

In cases where diagnostic attributes consist of salient attributes of a thing, we predict that even though implicitly trained contributors formulate inclusion rules themselves, they would report a similar number of diagnostic attributes, i.e., attributes that constitute inclusion rules, about an observed thing as the explicitly trained. In the same vein, untrained contributors are expected to apply a stimuli-driven approach to attention allocation without any prior knowledge of the diagnostic attributes (Itti and Koch 2001; Niebur and Koch 1996; Wolfe 1994). When most diagnostic attributes are salient, untrained contributors are therefore expected to also report similar numbers of diagnostic attributes as the implicitly and explicitly trained contributors. There will be no significant difference in the number of diagnostic attributes reported by untrained, implicitly trained, and explicitly trained contributors.

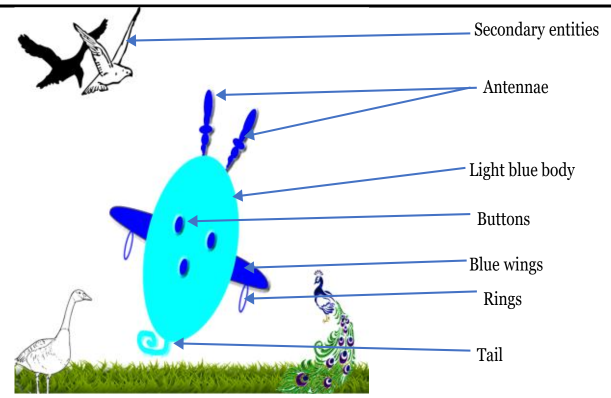
However, when the crowdsourcing task favors precision and classification accuracy, we expect explicitly trained contributors to be better at accurately identifying entities than implicitly trained contributors because explicitly trained contributors have learned clearly defined rules while implicitly trained contributors had to derive an inclusion rule by observing the salient attributes of every example they were presented with. Also, we expect that untrained contributors would not be able to classify the entity at the level of trained contributors. Untrained contributors do not know the specific class of the thing and would report basic (general) categories of the entity.

H3: Information quality. *Explicitly trained contributors will more accurately classify a primary thing than implicitly trained and untrained contributors, but all groups will report a similar number of diagnostic attributes of the primary thing*.

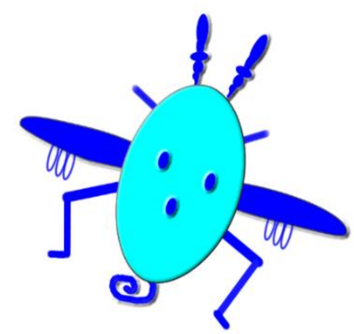
Experimental Design

We designed an experiment in the context of citizen science, a type of crowdsourcing in which citizens contribute to data collection and/or analysis while gaining scientific knowledge through their involvement in the research. Citizen science projects often seek knowledgeable contributors (Wiggins et al. 2011). Many citizen science projects are interested in discoveries (Lukyanenko, Parsons, et al., 2019). Citizen science is, therefore, a suitable context to test the impact of training on information diversity.

Following Kloos and Sloutsky's (2008) artificial stimuli, we created two classes of artificial insects: tyrans and nontyrans. We used novel artificial creatures to limit the effect of contributor prior knowledge on the study. The artificial insects are the primary entities of interest in our study. We defined tyran as a class (species) of artificial insects whose members meet a classification rule (a set of attributes and values of these attributes). Similar artificial stimuli that do not meet this classification rule are nontyrans. The classification rule consists of five requirements: *tyrans have (1) a short tail, (2) light blue bodies, (3) two or three buttons on their light blue bodies, (4) blue wings, and (5) either one or two rings on each blue wing*. The criteria for a primary entity to be classified as a nontyran is that it does not possess all the attributes stated in the classification rule. Figure 2 shows a sample tyran and a sample nontyran used in the experiment.



Tyran. It follows the classification rule.



Nontyran because it has three rings on each wing. The number of legs is not diagnostic.

Figure 2. Sample Tyran and Nontyran Images

The experiment consists of twenty images, each presented on a slide. Sixteen slides (the test images) are a mixture of tyrans and nontyrans. Four images containing catch items were placed intermittently among the test images to check whether the participants paid attention to the task. The catch items were differently shaped/colored stimuli that were not insects, and each participant was expected to report these stimuli correctly. The slides were presented in a nonrandomized order to all three groups, with image 5, 10, 15, and 20 containing catch items not related to the actual task. Each image has one primary entity but may not include secondary entities. Secondary entities presented in the images are everyday things such as birds, insects, and fences.

To test the reporting of attribute variations of the primary entities, six of the sixteen test images included variations of intrinsic attributes of primary entities relative to rules and training examples, i.e., 3 manipulations for tyrans and another 3 for nontyrans. For example, some primary entities had an increase in the number of a part of the thing (e.g., extra antennae), or different color parts, unlike the samples presented to participants during training.

We pretested the experimental materials with 12 ecology students who are familiar with observing and reporting organisms in the field. For the pretest, our goal was to examine the suitability of prompts to elicit unbiased responses from contributors. Also, we sought to test the ease with which implicit learners can decipher the inclusion rules from the samples they are presented during training, and the ease with which explicit learners can learn our classification rules. We tested participants using a 2-attribute inclusion rule and a 5-attribute inclusion rule. We also provided testers with stimuli with 2 attributes in common and stimuli with 5 attributes in common to see how well they decipher the inclusion rule. We displayed images of the entities in separate PowerPoint slides.

From the pretest, we found that asking contributors a nonleading question, i.e., "what do you see?"[3] was less biasing than asking them to identify the entity they observed or to report their sighting. Also, participants who used the 5-attribute inclusion rule did equally well as those who used the 2-attribute inclusion rule. However, in the feedback we received, participants in the 2-attribute inclusion rule condition complained that the 2-attribute inclusion rule made the task too easy. We conducted another pretest to examine the effect of the changes made based on our initial pretest.

Fifteen business students participated in the second pretest, which identified the appropriate duration of the reporting tasks for each image to be 40 seconds and confirmed the effectiveness of the changes made from the first pretest. These changes include the use of 5-attribute stimuli and the "what do you see?" prompt for the main experiment.

Task

We asked participants to imagine that designers of a game, similar to Pokémon Go, required their assistance in improving the design of an aspect of the game. The game requires players to observe artificial insects. Specific insects called Tyrans are harmful and can kill a player's character, while other insects called nontyrans, with some features similar to the tyrans but dissimilar in other features, can provide energy to a player's character in the game. The designers, therefore, needed to test if the participants can report data about these insects when observed to help improve their game. The participants were further informed that the goal of the experiment is to examine how people report the things they observe.

Participants were issued data entry booklets in which to write down their observations. For each image, each numbered page of the booklet required participants to record their observation about a correspondingly numbered image slide that we project. The prompt on each page of the booklet was, "what do you see?" Participants were also required to complete a demographic section after the study.

Participants

The 93 students who participated in this experiment were undergraduate students of Memorial University of Newfoundland. Students participated for course credit, for donations to their class graduation, and for the chance to win a bookstore gift card. After screening for completeness and the attentiveness of the contributor using the embedded catch items, responses from 84 participants, 28 participants per group was analyzed. Submissions from nine participants were excluded because of illegible writing, failure to report at least 3 out of 4 catch items correctly, and incomplete reports; 36 of the participants identified as male, and 48 as female.

We randomly assigned participants to three groups: (1) Explicitly Trained, (2) Implicitly Trained, and (3) Untrained. Members of the explicitly trained group were taught the classification rule introduced above for classifying the primary entities as tyrans or nontyrans. To increase their familiarity with the task, participants in the explicitly trained group were also shown five sample tyrans sufficient to decipher the classification rule, asked if they were tyrans, and given feedback on why these entities qualified as tyrans. We only showed participants images of tyrans because there are infinitely many ways the attributes of a primary entity may violate the classification rules (here we also follow Kloos and Sloutsky (2008)). We briefed participants in the implicitly trained group on the task they will perform and showed them the same five target stimuli used to teach the explicitly trained group, one at a time, to allow them to elicit classification criteria. The participants were allowed to study each image; however, we did not provide explicit rules to members of this group, nor did we give them feedback on their ability to determine whether a thing is a tyran or not. Members of the untrained group were not shown any sample images. However, like the other groups, members of the untrained group were informed that we were interested in examining how people report information.

[3] This is the prompt used by eBird, a popular citizen science platform (www.ebird.org).

Measures

We developed a coding scheme that accounts for attributes of the primary entities and attributes of other secondary entities reported by participants. Two of the authors participated in coding the first ten reports to establish consensus and conformance with the coding scheme. The first author coded the remaining reports, while the other authors reviewed the coded data at different stages of the coding process. The variables coded for are presented in Table 1.

Codes	Description
Classification Accuracy	The accuracy of identification of the primary entity as a tyran or nontyran
*Information Diversity	Information diversity is the total of all attributes reported about the entity. It is the sum of all intrinsic and mutual attributes of all the entities presented
Diagnostic Attributes	The number of primary entity diagnostic attributes reported
Intrinsic Attribute 1	The number of primary entity intrinsic attributes reported, i.e., attributes that depend on the primary entity only
Mutual Attribute 1	The number of primary entity mutual attributes (attributes that show an interaction between the primary entities and secondary entities)
Intrinsic Attribute 2	The number of secondary entity intrinsic attributes reported, i.e., attributes that depend on the secondary entity only
Mutual Attribute 2	The number of secondary entity mutual attributes (attributes that show an interaction between the secondary entities and the primary entity or other entities)
Attribute Variations	The number of reported new intrinsic attributes (attributes that vary from the ones presented in training)
Diagnostic Variations	The number of reported variations in intrinsic attributes of a primary entity that affect diagnostic parts of the entity. E.g., the wings are supposed to be blue. A new instance may have shorter than usual blue wings
Nondiagnostic Variations	The number of reported variations of nondiagnostic attributes (i.e., different from those seen in training). E.g., extra antennae. Antennae are not diagnostic
Secondary Entities	The number of secondary entities reported

*derived from coding intrinsic and mutual attributes of entities

Table 1. Variables Coded in the Contributed Data

Results

We used one-way analysis of variance (ANOVA) and Tukey's HSD[4] test for post-hoc comparison of the group averages (excluding the catch item images used for screening purposes only) to compare the variables described in Table 1 above, across the groups. We also used chi-square to test the difference in classification accuracy between the groups. We present the results of these analyses on the effect of training on our dependent variables: the reporting of diverse data, the reporting of attribute variations, and the reporting of quality data.

[4] Tukey's Honestly Significant Difference (Tukey's HSD) corrects for multiple comparisons (Homack, 2001)

Diverse Data

Information diversity is significantly different across groups, with $F(2,81) = 85.967$, $p < 0.001$. Table 2 shows that for information diversity, the group mean of the untrained group is significantly higher than those of the explicitly trained and implicitly trained groups. Also, the group average of the implicitly trained group is significantly greater than that of the explicitly trained group. This supports Hypothesis 1.

Information Diversity	A	B	mean(A)	mean(B)	Mean Diff.	Std. Err	T	p-value	η^2
	E	I	4.070	7.433	-3.363	0.383	-8.777	0.001	0.079
	E	U	4.070	8.984	-4.914	0.383	-12.825	0.001	0.155
	I	U	7.433	8.984	-1.551	0.383	-4.048	0.001	0.018

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 2. Results for the Effect of Training on Information Diversity

To better understand the results, we looked at the effect of training on the variables that make up the information diversity aggregate (Table 3).

	A	B	mean(A)	mean(B)	Mean Diff.	Std. Err	T	p-value	η^2
Intrinsic Attribute 1	E	I	2.217	4.621	-2.403	0.337	-7.136	0.0010	0.054
	E	U	2.217	3.404	-1.187	0.337	-3.524	0.0012	0.014
	I	U	4.621	3.404	1.217	0.337	3.612	0.0010	0.014
Mutual Attribute 1	E	I	0.725	0.748	-0.022	0.171	-0.131	0.900	0.000
	E	U	0.725	2.255	-1.529	0.171	-8.940	0.001	0.082
	I	U	0.748	2.255	-1.507	0.171	-8.810	0.001	0.080
Intrinsic Attribute 2	E	I	0.246	1.261	-1.016	0.162	-6.274	0.001	0.042
	E	U	0.246	1.730	-1.484	0.162	-9.170	0.001	0.086
	I	U	1.261	1.730	-0.469	0.162	-2.896	0.011	0.009
Mutual Attribute 2	E	I	0.882	0.804	0.078	0.164	0.477	0.8852	0.000
	E	U	0.882	1.596	-0.714	0.164	-4.357	0.0010	0.021
	I	U	0.804	1.596	-0.792	0.164	-4.834	0.0010	0.025

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 3. Results for the Effect of Training on Attribute Types

Table 3 shows that the mean value of the primary entities' intrinsic attributes (Intrinsic Attribute 1) is significantly different across the groups, with $F\ (2{,}1341) = 60.405$, $p = 0.001$. The average number of intrinsic attributes reported for the primary entity is highest for the implicitly trained group and lowest for the explicitly trained group. For the mutual attributes of the primary entities (Mutual Attributes 1), the group means of the explicitly trained and implicitly trained groups are significantly lower than that of the untrained group. However, there is no difference between the explicitly trained and implicitly trained groups. Untrained contributors were better than trained contributors at reporting the mutual attributes of the primary entities, but implicitly trained contributors reported more intrinsic attributes than explicitly trained and untrained contributors.

For images that contained secondary entities, the untrained group reported more intrinsic attributes of secondary entities (Intrinsic Attribute 2) than the trained contributors. The explicitly trained group reported the fewest intrinsic attributes. The mutual attributes of secondary entities (Mutual Attribute 2) were significantly different across groups. The mean for the untrained group was higher than those of the explicitly trained and implicitly trained groups; however, there is no significant difference between the explicitly trained and implicitly trained groups.

Attribute Variations

From the ANOVA results, the average number of attribute variations reported was different across the groups ($F(2{,}1341) = 8.964, p = 0.000$) at the 5% level of significance. The results of the post hoc tests presented in Table 4 also show that implicitly trained contributors do better than other groups at reporting variations in attributes, and there is no difference between the number of attribute variations reported by the untrained contributors and the explicitly trained contributors. This result support hypothesis 2.

	A	B	mean(A)	mean(B)	Mean Diff.	Std. Err	T	p-value	η^2
Attribute Variations	E	I	0.123	0.3795	-0.257	0.07	-3.666	0.001	0.015
	E	U	0.123	0.123	0.000	0.07	0.000	0.900	0.000
	I	U	0.380	0.123	0.257	0.07	3.666	0.001	0.015

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 4. Results for the Reporting of Attribute Variations

Breaking the number of variations in attributes reported for the observed instances of the primary entity into diagnostic attributes variations (Diagnostic Variations) and nondiagnostic attributes variations (Nondiagnostic Variations), we found that the groups were not significantly different at reporting attribute variations that affected classification. However, implicitly trained contributors were better at reporting attribute variations on nondiagnostic attributes, while untrained and explicitly trained contributors reported similar numbers of variations of nondiagnostic attributes. The results are presented in Table 5.

	A	B	mean(A)	mean(B)	Mean Diff.	Std. Err	T	p-value	η^2
Diagnostic Variations	E	I	0.112	0.101	0.011	0.044	0.252	0.900	0.000
	E	U	0.112	0.056	0.056	0.044	1.261	0.419	0.002
	I	U	0.101	0.056	0.045	0.044	1.008	0.664	0.001
Nondiagnostic Variations	E	I	0.011	0.280	-0.268	0.052	-5.150	0.001	0.029
	E	U	0.011	0.067	-0.056	0.052	-1.073	0.610	0.001
	I	U	0.280	0.067	0.212	0.052	4.077	0.001	0.018

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 5. Results for the Reporting of Variations of Diagnostic and Nondiagnostic Attributes

Quality Data

We tested for the effect of training on classification accuracy – the ability of contributors to correctly identify the primary entity as a tyran or a nontyran. We expected that untrained contributors would not be able to classify the entity and would instead report its basic categories, e.g., bug or insect, and describe the entity by its attributes. We, therefore, excluded untrained contributors from the test for classification accuracy.

Identification or classification accuracy is valued as one or zero for each target entity presented, depending on whether the contributor correctly identifies the target entity as either a tyran or a nontyran (1) or not (0). The result is presented in Table 6 below.

Classification Accuracy			
Group	Incorrect	Correct	Total
E	62 (13.8%)	386 (86.2%)	448(100%)
I	162(36.2%)	286(63.8%)	448(100%)
Total	224	672	895

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 6. Differences in Accuracy for Explicitly and Implicitly Trained Groups

Following the result of the classification accuracy test, we then compared the extent to which trained and untrained contributors reported diagnostic attributes of the target entity. We found that the number of diagnostic attributes reported is not significantly different across groups, with $F(2,1341) = 0.92$, $p = 0.399$ at the 5% level of significance. Post-hoc analysis also reveals significant similarity between all the number of diagnostic attributes reported by all groups (see Table 7).

	A	B	mean(A)	mean(B)	Mean Diff.	Std. Err	T	p-value	η^2
Diagnostic Attributes	E	I	1.514	1.440	0.074	0.183	0.407	0.900	0.000
	E	U	1.514	1.272	0.242	0.183	1.324	0.383	0.002
	I	U	1.440	1.272	0.167	0.183	0.917	0.718	0.001

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 7. Results for the Reporting of Diagnostic Attributes (Completeness)

These results support hypothesis 3.

Manipulation Check

To confirm that the trained contributors exhibited selective attention to the primary entity and its attributes, and a learned inattention to other entities, we compared their reporting of the presence of secondary entities. Since secondary entities are familiar entities, we examined the degree to which each group reported the presence of these entities as proof for whether they paid selective attention to primary entities. Applying One-way ANOVA to the data of images that included secondary entities, we found that untrained contributors reported the occurrence of secondary entities more often than trained contributors (Table 8).

	A	B	mean(A)	mean(B)	Mean Diff.	Std. Err	T	p-value	η^2
Secondary Entities	E	I	1.036	1.544	−0.508	0.109	−4.674	0.001	0.029
	E	U	1.036	2.289	−1.253	0.109	−11.521	0.001	0.154
	I	U	1.544	2.289	−0.745	0.109	−6.847	0.001	0.060

E=Explicitly trained group, I= Implicitly trained group, U= Untrained group

Table 8. Differences in the Reporting of Secondary Entities

Discussion

Contributors to observational crowdsourcing projects are frequently trained based on the assumption that knowledgeable contributors will report better data than untrained contributors. However, our study raises important questions about the validity of this assumption. We found through our experiment that training affects contributors' ability to report diverse data. Untrained contributors report more diverse data than trained contributors. Trained and untrained contributors were equally capable of reporting variations to diagnostic attributes (the attributes essential for classification). However, implicitly trained contributors reported more variations to nondiagnostic attributes than explicitly trained and untrained contributors. Also, both trained and untrained contributors can report diagnostic attributes that will lead to accurate classification, but trained contributors, particularly explicitly trained contributors, are better at accurately classifying things (see the summary of results in Table 9).

Our results suggest that untrained contributors might be best suited to such situations where data needs to be collected about fleeting phenomena, and data users want to capture as much detail (i.e., diversity of perspectives) as they can the first time. Also, using untrained contributors will not lead to loss of information quality when data users can use machine learning (Lukyanenko et al. 2019) or even a second layer of analyses by experts (Brynjolfsson et al. 2015; Lukyanenko et al. 2019) to determine entities from reported attributes. Restricting participation to knowledgeable contributors or insisting on training contributors is unnecessary in many observational crowdsourcing cases and would be limiting the inclusiveness of projects, potentially depriving both participants and data users of knowledge that emerges from a crowdsourcing endeavor.

Furthermore, diagnostic attribute variations can indicate a new variant of an entity, a new species, or a new stage of the development of an entity. All these may portend useful insight that can lead to discoveries about a thing (Mahner and Bunge 1997). If the goal of an observational crowdsourcing task includes making discoveries by collecting reports of variations in diagnostic attributes, then training offers no benefits to observational crowdsourcing projects. However, data users may also be interested in nondiagnostic attribute variations, and in such instances, it makes sense to train contributors implicitly but not explicitly. Overall, our results suggest no need to train crowds because of a concern for data quality. Given that repurposable data are more variable than homogeneous, streamlined data (Günther et al. 2017), training contributors negatively impact their ability to provide data that can be flexibly used in new contexts to drive insightful decision-making (Ogunseye & Parsons 2018).

This research has broad implications for crowdsourcing systems that explicitly train their contributors and crowdsourcing systems that inadvertently train crowds by allowing the same contributors to participate in the same task over and over again – the business model of crowd-hiring platforms like Amazon Mechanical Turk (Ogunseye et al. 2017). It can serve as a guide to system designers and decision-makers, helping to ensure that their decision to train contributors aligns with the goals of their projects. The research also contributes to theory by highlighting how indiscriminate crowd-reuse in online review systems and crowdsourcing platforms can affect the value of data collected through these systems.

At the same time, this study answers the call for more research that guide the designs of information systems that augment human intelligence and mitigate human limitations (Jain et al. 2018). The designs of observational crowdsourcing systems, whether for tracking disease outbreaks like COVID 19 and SARS or collecting information about flora and fauna in a region of North America, need to mitigate the human tendency to focus on known uses of data and potentially leave out valuable information. This study can provide prescriptive guidance to practitioners and researchers on how to compose their crowds or design their crowd-facing systems to mitigate the negative consequences of learning on the data they collect. In the psychology and cognitive science domains, this research increases our knowledge of how selective attention affects information people produce and not just the information they consume.

There are several limitations to the generalizability of our findings. First, in our experiment, we used five exemplars in the implicit training condition. The number of exemplars may not be adequate to learn all the rules for classifying entities and may be too many in some cases. Second, we assume organizations and individuals that own and use observational crowdsourcing platforms can classify entities from collected data containing attributes and basic classes. However, this might not be easy or realistic in every case. In this study, we can understand the descriptions provided by untrained contributors about the entities, and the images used in the experiment are accessible to us for confirmation. In some real-world scenarios, data users might not understand descriptions provided by untrained contributors in the field.

Hypotheses	**Comments on Findings**	**Supported**
H1: Information diversity. *Untrained contributors>implicitly trained contributors,>explicitly trained contributors.*	Untrained contributors reported more diverse data than trained contributors.	Yes
H2: Attribute variations. *Explicitly trained contributors < untrained contributors <implicitly trained contributors*	When the varying attributes are diagnostic, *untrained ≡ trained. For all variations,* implicitly trained contributors > explicitly trained and untrained contributors.	Partially
H3: Information quality. *(classification accuracy) Explicitly trained contributors > implicitly trained > untrained contributors; (completeness) untrained ≡ trained*	True in all cases	Yes

Table 9: Summary of Hypotheses and Findings

Conclusion

Users of observational crowdsourcing who train their crowds believe trained contributors will provide better quality data than untrained contributors. Even though training may increase classification accuracy in crowdsourced data, this research shows that training reduces data diversity, with potentially negative consequences for the repurposability of collected data.

Both untrained and trained contributors can adequately report variations to diagnostic attributes of different instances of a thing, which may lead to discoveries. However, if we must train our crowds, it may be better to employ implicit training than explicit training. Nonetheless, there are scenarios where precision and classification are of absolute necessity, and at these times, explicit rule-based training would be ideal.

Moreover, data requirements can change at several stages of a decision-making process - during data collection or even after the initial analytics results come in. Because we do not always know precisely how our crowdsourced data will be used, and if the information we require from data will change, there is a need to collect data that can accommodate known and emergent uses. The result of this study suggests that the choice to train our crowds would, therefore, depend on whether we need repurposable data. Other factors we need to consider are how much value we place on the inclusiveness of our project and whether the data that untrained contributors can provide would be usable in the specific context of our projects. But knowledgeable contributors are not required to acquire quality data.

Further study is necessary to test the effect of selective attention in the field to aid our research's generalizability. It would also be useful to understand how to mitigate the impact of selective attention on collected data through task and system designs. Beyond the crowdsourcing space, we will continue to work on guiding the collection of repurposable data.

References

Berry, D., and Dienes, Z. 1993. Towards a Characterization of Implicit Learning.

Best, C. A., Yim, H., and Sloutsky, V. M. 2013. "The Cost of Selective Attention in Category Learning: Developmental Differences between Adults and Infants," Journal of Experimental Child Psychology (116:2), pp. 105–119.

Bjorklund, D. F., and Harnishfeger, K. K. 1990. "The Resources Construct in Cognitive Development: Diverse Sources of Evidence and a Theory of Inefficient Inhibition," Developmental Review (10:1), Elsevier, pp. 48–71.

Bonter, D. N., and Cooper, C. B. 2012. "A Process for Improving Data Quality and a Strategy for Ensuring Sustainability in a Citizen Science Project," Front Ecol Envir (10), pp. 305–307.

Brynjolfsson, E., Geva, T., and Reichman, S. 2015. "Crowd-Squared: Amplifying the Predictive Power of Search Trend Data," Brynjolfsson, E., Geva, T., & Reichman, S., Crowd-Squared: Amplifying the Predictive Power of Search Trend Data. MIS Quarterly (Forthcoming).

Bunge, M. 1977. Treatise on Basic Philosophy: Ontology I: The Furniture of the World, (Vol. 3), Springer Science & Business Media.

Buschman, T. J., and Miller, E. K. 2007. "Top-down versus Bottom-up Control of Attention in the Prefrontal and Posterior Parietal Cortices," Science (315:5820), pp. 1860–1862.

Gazzaley, A., and Nobre, A. C. 2012. "Top-down Modulation: Bridging Selective Attention and Working Memory," Trends in Cognitive Sciences (16:2), pp. 129–135.

Gelman, S. A., and Markman, E. M. 1986. "Categories and Induction in Young Children," Cognition (23:3), pp. 183–209.

Ghasemaghaei, M., and Calic, G. 2019. "Can Big Data Improve Firm Decision Quality? The Role of Data Quality and Data Diagnosticity," Decision Support Systems (120), pp. 38–49.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., and Feldberg, F. 2017. "Debating Big Data: A Literature Review on Realizing Value from Big Data," The Journal of Strategic Information Systems.

"Hanny's Voorwerp – History of a Mystery." 2013. Daily Zooniverse, September 24. (https://daily.zooniverse.org/2013/09/24/hannys-voorwerp/, accessed April 25, 2020).

Harrison, C., Jorder, M., Stern, H., Stavinsky, F., Reddy, V., Hanson, H., Waechter, H., Lowe, L., Gravano, L., and Balter, S. 2014. "Using Online Reviews by Restaurant Patrons to Identify Unreported Cases of Foodborne Illness—New York City, 2012–2013," MMWR. Morbidity and Mortality Weekly Report (63:20), p. 441.

Hoffman, A. B., and Rehder, B. 2010. "The Costs of Supervised Classification: The Effect of Learning Task on Conceptual Flexibility.," Journal of Experimental Psychology: General (139:2), p. 319.

Itti, L., and Koch, C. 2001. "Computational Modelling of Visual Attention," Nature Reviews Neuroscience (2:3), p. 194.

Jain, H., Padmanabhan, B., Pavlou, P. A., and Santanam, R. T. 2018. "Call for Papers—Special Issue of Information Systems Research—Humans, Algorithms, and Augmented Intelligence: The Future of Work, Organizations, and Society," Information Systems Research (29:1), INFORMS, pp. 250–251.

Jensen, M. L., Dinger, M., Wright, R. T., and Thatcher, J. B. 2017. "Training to Mitigate Phishing Attacks Using Mindfulness Techniques," Journal of Management Information Systems (34:2), Taylor & Francis, pp. 597–626.

Johnston, W. A., and Dark, V. J. 1986. "Selective Attention," Annual Review of Psychology (37:1), pp. 43–75.

Katsuki, F., and Constantinidis, C. 2014. "Bottom-Up and Top-Down Attention: Different Processes and Overlapping Neural Systems," The Neuroscientist (20:5), pp. 509–521.

Keil, F. C. 1989. Concepts, Kinds, and Conceptual Development, Cambridge, MA: MIT Press.

Kiwelekar, A. W., and Joshi, R. K. 2010. "An Object-Oriented Metamodel for Bunge-Wand-Weber Ontology," ArXiv Preprint ArXiv:1004.3640.

Kloos, H., and Sloutsky, V. M. 2008. "What's behind Different Kinds of Kinds: Effects of Statistical Density on Learning and Representation of Categories.," Journal of Experimental Psychology: General (137:1), p. 52.

Leonardi, P. M., and Bailey, D. E. 2008. "Transformational Technologies and the Creation of New Work Practices: Making Implicit Knowledge Explicit in Task-Based Offshoring," MIS Quarterly, JSTOR, pp. 411–436.

Lukyanenko, R., and Parsons, J. 2019. "Beyond Micro-Tasks: Research Opportunities in Observational Crowdsourcing," in Crowdsourcing: Concepts, Methodologies, Tools, and Applications, IGI Global, pp. 1510–1535.

Lukyanenko, R., Parsons, J., Wiersma, Y. F., and Maddah, M. 2019. "Expecting the Unexpected: Effects of Data Collection Design Choices on the Quality of Crowdsourced User-Generated Content," MIS Quarterly (43:2), pp. 623–647.

Mahner, M., and Bunge, M. 1997. Foundations of Biophilosophy, Springer Science & Business Media.

Neill, W. T., Valdes, L. A., and Terry, K. M. 1995. "7 - Selective Attention and the Inhibitory Control of Cognition," in Interference and Inhibition in Cognition, F. N. Dempster, C. J. Brainerd, and C. J. Brainerd (eds.), San Diego: Academic Press, pp. 207–261.

Nelson, R. R., Todd, P. A., and Wixom, B. H. 2005. "Antecedents of Information and System Quality: An Empirical Examination within the Context of Data Warehousing," Journal of Management Information Systems (21:4), pp. 199–235.

Niebur, E., and Koch, C. 1996. "Control of Selective Visual Attention: Modeling the" Where" Pathway," in Advances in Neural Information Processing Systems, pp. 802–808.

Nsoesie, E. O., Kluberg, S. A., and Brownstein, J. S. 2014. "Online Reports of Foodborne Illness Capture Foods Implicated in Official Foodborne Outbreak Reports," Preventive Medicine (67), pp. 264–269.

Ogunseye, S., and Parsons, J. 2016. "Can Expertise Impair the Quality of Crowdsourced Data?," SIGOPEN Developmental Workshop at ICIS 2016.

Ogunseye, S., and Parsons, J. 2018. "Designing for Information Quality in the Era of Repurposable Crowdsourced User-Generated Content," in International Conference on Advanced Information Systems Engineering, Springer, pp. 180–185.

Ogunseye, S., Parsons, J., and Lukyanenko, R. 2017. Do Crowds Go Stale? Exploring the Effects of Crowd Reuse on Data Diversity, WITS.

Palmer, J. R. B., Oltra, A., Collantes, F., Delgado, J. A., Lucientes, J., Delacour, S., Bengoa, M., Eritja, R., and Bartumeus, F. 2017. "Citizen Science Provides a Reliable and Scalable Tool to Track Disease-Carrying Mosquitoes," Nature Communications (8). (https://doi.org/10.1038/s41467-017-00914-9).

Parsons, J., and Wand, Y. 2000. "Emancipating Instances from the Tyranny of Classes in Information Modeling," ACM Transactions on Database Systems (TODS) (25:2), pp. 228–268.

Parsons, J., and Wand, Y. 2014. "A Foundation for Open Information Environments," Proceedings of the European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel, June 9-11, 2014

Plude, D. J., Enns, J. T., and Brodeur, D. 1994. "The Development of Selective Attention: A Life-Span Overview," Acta Psychologica (86:2), pp. 227–272. (https://doi.org/10.1016/0001-6918(94)90004-3).

Prat-Ortega, G., and de la Rocha, J. 2018. "Selective Attention: A Plausible Mechanism Underlying Confirmation Bias," Current Biology (28:19), Elsevier, pp. R1151–R1154.

Ransbotham, S., and Kiron, D. 2017. "Analytics as a Source of Business Innovation," MIT Sloan Management Review; Cambridge (58:3).

Reber, A. S. 1993. Oxford Psychology Series, No. 19. Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious, Oxford University Press.

Ritchie, S. A., Long, S., Hart, A., Webb, C. E., and Russell, R. C. 2003. "An Adulticidal Sticky Ovitrap for Sampling Container-Breeding Mosquitoes," Journal of the American Mosquito Control Association (19:3), [Fresno, Calif.]: The Association,[c1985-, pp. 235–242.

Rosch, E. H. 1973. "Natural Categories," Cognitive Psychology (4:3), pp. 328–350.

Rosemann, M., and Green, P. 2002. "Developing a Meta Model for the Bunge–Wand–Weber Ontological Constructs," Information Systems (27:2), pp. 75–91.

Rozendaal, J. A., Staff, W. H. O., santé, O. mondiale de la, and Organization, W. H. 1997. Vector Control: Methods for Use by Individuals and Communities, World Health Organization.

Sadiq, S., and Indulska, M. 2017. "Open Data: Quality over Quantity," International Journal of Information Management (37:3), pp. 150–154.

Schomberg, J. P., Haimson, O. L., Hayes, G. R., and Anton-Culver, H. 2016. "Supplementing Public Health Inspection via Social Media," PloS One (11:3), p. e0152117.

Shanks, G., Tansley, E., Nuredini, J., Tobin, D., and Weber, R. 2008. "Representing Part-Whole Relations in Conceptual Modeling: An Empirical Evaluation," MIS Quarterly, pp. 553–573.

Simons, D. J., and Ambinder, M. S. 2005. "Change Blindness: Theory and Consequences," Current Directions in Psychological Science (14:1), pp. 44–48.

Simons, D. J., and Rensink, R. A. 2005. "Change Blindness: Past, Present, and Future," Trends in Cognitive Sciences (9:1), pp. 16–20.

Sprenger, P. R. D. 1987. "The Used Tire Trade: A Mechanism for the Worldwide Dispersal of Container Breeding Mosquitoes," J. Am. Mosq. Control. Assoc (3), p. 494.

Sun, R., Slusarz, P., and Terry, C. 2005. "The Interaction of the Explicit and the Implicit in Skill Learning: A Dual-Process Approach.," Psychological Review (112:1), American Psychological Association, p. 159.

Surowiecki, J. 2005. The Wisdom of Crowds, Anchor.

Taylor, H. 2007. "Tacit Knowledge: Conceptualizations and Operationalizations," International Journal of Knowledge Management (IJKM) (3:3), IGI Global, pp. 60–73.

Theeuwes, J. 2010. "Top-down and Bottom-up Control of Visual Selection," Acta Psychologica (135:2), pp. 77–99.

Vargios, E. M. 2007. Educational Psychology Research Focus, Nova Publishers.

Wand, Y., Storey, V. C., and Weber, R. 1999. "An Ontological Analysis of the Relationship Construct in Conceptual Modeling," ACM Trans. Database Syst. (24:4), pp. 494–528.

Wand, Y., and Wang, R. Y. 1996. "Anchoring Data Quality Dimensions in Ontological Foundations," Communications of the ACM (39:11), pp. 86–96.

Wand, Y., and Weber, R. 1990. "Mario Bunge's Ontology as a Formal Foundation for Information Systems Concepts," Studies on Mario Bunge's Treatise (18).

Wang, L., Gopal, R., Shankar, R., and Pancras, J. 2015. "On the Brink: Predicting Business Failure with Mobile Location-Based Checkins," Decision Support Systems (76), pp. 3–13.

Wang, R., and Strong, D. M. 1996. "Beyond Accuracy: What Data Quality Means to Data Consumers," Journal of Management Information Systems (12:4), pp. 5–33.

Wiggins, A., and He, Y. 2016. "Community-Based Data Validation Practices in Citizen Science," in Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW (Vol. 27), Association for Computing Machinery, pp. 1548–1559. (https://doi.org/10.1145/2818048.2820063).

Wiggins, A., Newman, G., Stevenson, R. D., and Crowston, K. 2011. "Mechanisms for Data Quality and Validation in Citizen Science," in 2011 IEEE Seventh International Conference on E-Science Workshops, December, pp. 14–19.

Wixom, B. H., and Todd, P. A. 2005. "A Theoretical Integration of User Satisfaction and Technology Acceptance," Information Systems Research (16:1), pp. 85–102.

Wolfe, J. M. 1994. "Guided Search 2.0 a Revised Model of Visual Search," Psychonomic Bulletin & Review (1:2), pp. 202–238.

Woodall, P. 2017. "The Data Repurposing Challenge: New Pressures from Data Analytics," Journal of Data and Information Quality (JDIQ) (8:3–4), p. 11.

Yang, R., Xue, Y., and Gomes, C. 2018. Pedagogical Value-Aligned Crowdsourcing: Inspiring the Wisdom of Crowds via Interactive Teaching.

"Zika, Mosquitoes, and Standing Water | | Blogs | CDC." 2016. Zika, Mosquitoes, and Standing Water, March. (https://blogs.cdc.gov/publichealthmatters/2016/03/zikaandwater/, accessed May 1, 2020).