Training citizen scientists: A qualitative, comparative, multiple case study to identify theoretical and instructional design themes in current citizen science training initiatives

Sigma Xi Student Research Showcase
PhD Education- Curriculum, Instruction, and Assessment
University of the Rockies
M. Gaddis
3/24/2018
Citizen scientists are volunteers who participate in scientific activities under the guidance of professional scientists and organizations.

This dissertation inquiry investigates citizen science training and its perceived effectiveness.

One of the first and longest running citizen science projects started by Thoreau at Walden Pond involves measuring the day of plant emergence in the spring.
The general problem is data collected by citizen scientists is often viewed as unreliable by the scientists and land managers who might use it, which results in lower confidence in and scientific use of citizen scientist-collected data.

The specific problem is the absence of educational training measurement in citizen science program design and analysis with which to ascertain the learning gains of trained citizen scientists.
The purpose of this qualitative comparative multiple case study is to identify patterns and themes in content, instructional design, theoretical alignment, and perceived efficacy of training for citizen scientists tasked with collecting ecological data in the field.
Citizen scientists are increasingly important to the implementation and on-going assessment of ecological restoration, species identification, ecological monitoring on natural lands (Bonney et al., 2014; Dickinson, Zuckerberg, & Bonter, 2010; Handel, Saito, & Takeuchi, 2013; Havlick et al., 2014; Maschinski, Wright, & Lewis, 2012; Schmeller et al., 2008), and myriad other scientific contexts.

Citizen scientists collect data to monitor plant populations.
The purpose of this research design is to identify patterns and themes in content, instructional design, theoretical alignment, and perceived efficacy of training for citizen scientist volunteers tasked with collecting data in the field.

- R1: What are the characteristics of citizen science trainings designed to train volunteers to collect ecological data in natural land settings?

- R2: How do organizational leaders describe their perception of the efficacy of the trainings to produce reliable data collection?
The literature currently focuses on three components of citizen science:

• science literacy gains for volunteers who participate in citizen science
  (Crall et al., 2013; Havlick et al., 2014; Lynch, 2016; Rasmussen, 2015)

• credibility of data collected by citizen scientists
  (reviewed in Dickinson, Zuckerberg, & Bonter, 2010; Kremen, Ullman, & Thorp, 2010; Reynolds, 2016; Storey & Wright-Stow, 2017)

• the usefulness and economic benefits of volunteers to natural resource management and monitoring
  (Gollan, De Bryun, Reid, & Wilkie, 2012; Handel et al., 2013)

Figure 1. ProQuest literature returns for citizen science with no date restrictions
Citizen science engages scientific researchers and:
- indigenous human populations  
  (Baker, 2013; Dolrenry, Hazzah, & Frank, 2016)
- college students  
  (Davis, Belote, Williamson, Larson, & Esch, 2016; Oberhauser & LeBuhn, 2012)
- students of all ages  
  (Doran & Montmerle, 2012)
- private land owners  
  (Dickinson, Zuckerberg, & Bonter, 2010)
- non-profit groups  
  (Clarridge, 2016)

Citizen science occurs in:
- forests  
  (Danielsen et al., 2014; Davis, Belote, Williamson, Larson, & Esch, 2016; Toman & Shindler, 2006)
- gardens  
  (Birkin & Goulson, 2015)
- oceans  
  (Koss et al., 2009; Van der Velde et al., 2017)
- rivers  
  (Kruse, 2014; Storey & Wright-Stow, 2017)
- the sky  
  (Doran & Montmerle, 2012; Henden, 2011)
- from within one’s own home or office with the use of computer-accessible cloud-based data  
  (Ricci, 2015)
• 5% of the citizen science literature references training
• Over 150 scholarly documents analyzed for this literature review
• The journal *Citizen Science: Theory and Practice* is the first peer-reviewed journal dedicated to reporting citizen science-related research
  - first issue published in 2016
  - 6 articles from this journal are cited in this literature review (2 addressed training)
• 10% of all articles in *Citizen Science: Theory and Practice* addressed training
Science Literacy Gains as a Result of Participation - slide 8

- Despite many peer-reviewed publications on the topic, citizen science projects are rarely started, if ever, for the sole purpose of enhancing science literacy.

Photo credit: http://www.openscientist.org/2013/01/the-levels-of-citizen-science.html
Citizen science has the potential to increase the scale of data collection, but data error and bias are poorly understood (Dickinson et al., 2010)

Discussions of data quality are the most numerous discussions in the literature currently (Clarridge, 2016; Crall et al., 2011; Danielsen et al., 2014; Gollan, De Bryun, Reid, & Wilkie, 2012; Jordan, Sorenson, & Ladeau, 2017; Kremen, Ullman, & Thorp, 2010; Koss et al., 2009; Moyer-Homer, Smith, & Belt, 2012; Reynolds, 2016; Schmeller et al., 2008; Storey & Wright-Stow, 2017)
• Training and learning in citizen science have not been reported in the literature or been evaluated as rigorously as the direct comparisons of data collection efforts

• A collection of papers support a simple conclusion that training is valuable regardless of its design
  (Gallo and Waitt, 2011, Fuccillo, Crimmins, De Rivera, and Elder, 2015), Van Horn, Zug, LaCombe, Velez-Liendo, & Paisley, 2014)
All studies that address training do so in one of two ways:
- a simple fact of programming
- training is the premise of the research

• Crall and Newman (numerous publications) found that citizen scientists
  • are proficient at collecting simple data, but their data quality decreases with increasing data collection challenge
  • perform equally well with live or video training, and better than static documents alone
  • struggle with procedural and GPS skills

• Starr et al. (2014) found that citizen scientists
  • perform equally well with live or video training, and better than static documents alone
• Educational experiences contribute to learning

• The acquisition of procedural knowledge is possible in a training setting outside of academia where scientists are formally trained

• Andragogy and backwards design are learning and design principles that are successful in other training contexts

  (Mitchell & Sonora, 2012; Salas et al., 2012; Smith, 2015; Toman & Shindler, 2006)
• A qualitative multiple case study design
  • appropriate for this inquiry because it affords a breadth of information from which themes can be extrapolated
  • no observational study currently exists to characterize citizen science training yet scholars are debating the quality of the data collected by citizen scientists (Danielsen et al., 2014; Smith, 2015)

• The research design is exploratory in nature, but the sequential investigative steps drive the investigation into the realm of descriptive case study design (Yin, 2013)
Methodology - slide 15

• Involves the collection of both text-based and numeric data into one analytical tool, a case study database managed with dedoose™ software.

• Sequential investigation
  • case identification
  • training document analysis
  • organizational survey
  • follow-up semi-structured interviews with training leaders from categorically-representative organizations
Qualitative Multiple Case Study Design - slide 16

- Involves the collection of both text-based and numeric data into one analytical tool, a case study database managed with dedoose™ software.

- Sequential investigation
  - case identification
  - training document analysis
  - organizational survey
  - follow-up semi-structured interviews with training leaders from categorically-representative organizations

![Diagram showing the process flow from data collection to triangulation of qualitative data.]

Data collected in the early phases will inform the semi-structured interviews as the interviews are intended to corroborate and/or clarify training themes already identified.

Triangulation of qualitative data promotes confidence in identified themes.
• Goal for Data Saturation: 30-50 cases in the case study database (Bertaux, 1981; Mason, 2010)

• If more than 100 organizations meet the criteria in case sample selection, more detailed selection criteria may determined

• The case study database will be used to identify common themes across cases
  • the document analysis will inform the development of the codes in the case study database
  • the survey will constitute the primary data collection tool
  • the semi-structured interviews will triangulate the document analysis and survey
Trustworthiness Concerns in Qualitative Studies - slide 18

• Collecting multiple streams of analogous data
  • facilitates the establishment of qualitative data quality
  • creates a redundant inquiry system in which codes and themes are revealed and confirmed or contradicted
    • Confirmation - demonstrates data credibility
    • Contradiction - the theme may be probed with additional inquiry

• When independent researchers arrive at similar codes and themes in independent analysis
  • confirmability is demonstrated

• When these codes and themes are applicable to a number of cases
  • transferability and dependability are demonstrated
Generalization of the Study - slide 19

• Scholarly information about citizen science training is limited to
  • single-organization case studies
  • training simulations

• This is the first scholarly attempt to aggregate training information

• Aim is to reveal similarities and differences among cases

Anyone lacking a terminal degree in science, including students, are citizen scientists.
• Qualitative, comparative, multiple case study design
• Purposive sample - CitSci.org database
• Sequential design
• Training document analysis
• Survey
• Semi-structured interviews
• Data analyzed using dedoose© software
• Data analysis process may triangulate the three data components of each case
• Cross-case analysis
  • reveal themes across citizen science training programs
  • reveal demographic and program trends across the citizen science landscape
• resulting conclusions may inform future citizen science training initiatives and citizen science data collection reliability


References - slide 22


References - slide 23


References - slide 24


• Thank you to my dissertation chair, Dr. Kim Fonteix, who has offered unwavering support throughout the process. Thank you to my committee, Dr. Greg Newman and Dr. Chuck Dull. Thank you to my family who made the biggest sacrifice when I decided to go back for my PhD while working and raising my three young children.