

# Secret Box: Augmenting Big Data with Local Data

---

Rachel Atherton | @wrathertweets | ratherto@purdue.edu

Data can be **dehumanizing** (Eubanks, 2018). Divorced from context, **big data is easily manipulated** to back up stories data collectors or purchasers want to tell (Knafllic, 2015), and algorithms that use big data sets often reproduce or reinforce systems of marginalization for already vulnerable populations (Noble, 2017; O'Neill, 2016). Quantitative data's ostensible objectivity obscures these realities, so we must continue bringing to light the material consequences of relying only on big data. In **non-profit settings** specifically, funders' requirements for data collection can negatively impact organizations' efficacy and even cause "data drift," which may cause organizations to pivot away from their missions – but researchers have shown that funders often prefer easily digestible data narratives and "feel-good quotes" to nuanced analyses (Bopp, Harmon, & Volda, 2017).

Food Finders, a medium-sized non-profit food bank working in Tippecanoe county and the surrounding area, collects a great deal of data about visitors; Food Finders' ability to obtain and maintain funding sources depends on the data they collect being consistent with funders' expectations, but that data isn't always what is most beneficial for clients to share. Based on a pilot study conducted with two other graduate students (Elizabeth Geib and Eliza Gellis) in Fall 2018, where we placed a "secret box" (Punch, 2002) in Food Finders' main pantry waiting area and invited clients to write, I ask:

**How can Food Finders leverage local knowledges through methods such as a secret box to address inaccuracies or inconsistencies in "objective" quantitative datasets?**

**The secret box works because it:**

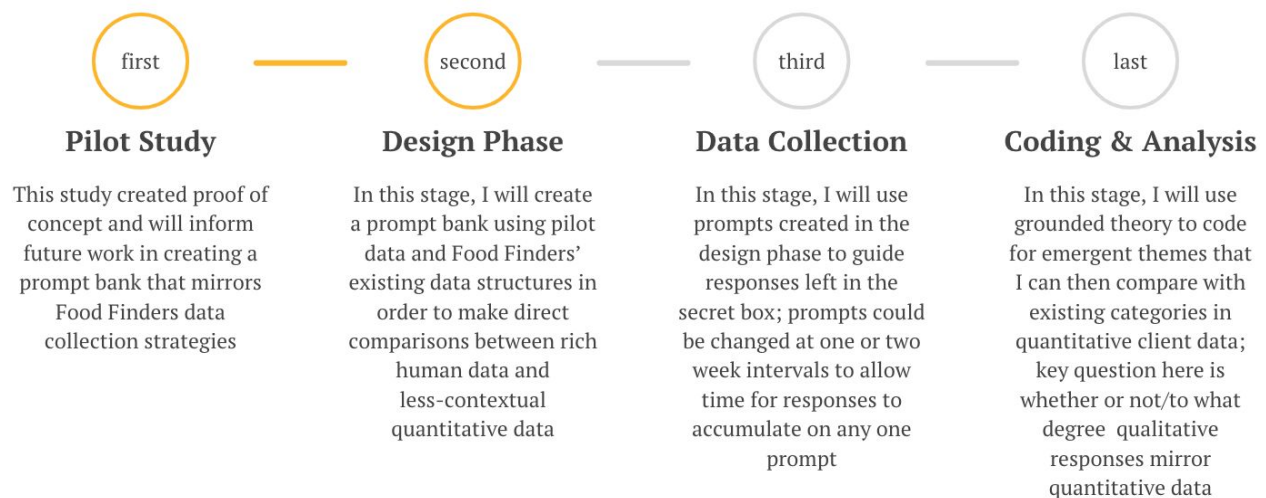
- Offers anonymous space to share thoughts and feelings
- Avoids barriers to entry as with digital surveys
- Leaves almost total agency with participants, since box is anonymous and prompt is optional

**Other secret box users include:**

- Punch (2002), where secret box augmented focus groups and various kinds of task-based interviews with adolescents ages 13-14
  - Emphasized that secret box helped mitigate uneven power distributions between young participants and adult researchers
  - Anonymity allowed participants to express thoughts and feelings they were unable or unwilling to share in group interviews or out loud where their responses would be associated with them (experiences of sexual abuse, for instance)
  - Suggested secret boxes as part of a larger group of methods in studies where power dynamics are unequal (like working with vulnerable populations at Food Finders, for instance)
- Lyndon (2018) based her work on Punch (2002).
  - Participants were adults, unlike Punch, but power dynamics still existed in focus groups, and secret box was used to allow participants to share thoughts they were unwilling to share in front of other

## Projected takeaways include:

- Since data is often dehumanizing, agency-denying, and marginalizing, especially to already vulnerable populations, being mindful of how we collect data from participants in settings such as non-profits is crucial.
- Methods such as the secret box, which are anonymous, have a low barrier to entry (as compared to digital surveys, for instance), and suggest but do not require a given prompt can empower participants to give data they want to share and hold onto data they don't.
- Non-profits usually cannot avoid participating in big data collection, but using qualitative methods to collect thick, rich data from individual clients can help augment fragmented quantitative data and avoid "data drift" and other negative impacts
- Ultimately, data that clients have more power over is data that helps nonprofits better serve clients.



## References & Further Reading

- Bopp, C., Harmon, E., & Volda, A. (2017). Disempowered by data: Nonprofits, social enterprises, and the consequences of data-driven work. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, 3608-3619.
- Bornakke, T., & Due, B. (2018). Big-thick blending: A method for mixing analytical insights from big and thick data sources. *Big Data & Society*, 5(1), 1-16.
- boyd, d., and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679.
- Eubanks, V. (2017). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York, NY: St. Martin's Press.
- Harris, C., Jackson, Mayblin, Piekut, & Valentine. (2015). 'Big Brother welcomes you': Exploring innovative methods for research with children and young people outside of the home and school environments. *Qualitative Research*, 15(5), 583-599.
- Knaflic, C. (2015). *Storytelling with data: A data visualization guide for business professionals*. Hoboken, NJ: Wiley.
- Lyndon, S. (2018). Analyzing focus groups about poverty in the early years using a narrative approach. *SAGE Research Methods Cases*.
- Merry, S. E. (2016). *The seductions of quantification: Measuring human rights, gender violence, and sex trafficking*. Chicago, IL: U of Chicago P.
- Moore, K., & Elliott, T. (2016). From participatory design to a listening infrastructure: A case of urban planning and participation. *Journal of Business and Technical Communication*, 30(1), 59-84.
- Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York, NY: New York University Press.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York, NY: Crown.
- Punch, S. (2002). Interviewing strategies with young people: The 'secret box,' stimulus material and task-based activities. *Children & Society*, 16(1), 45-56.