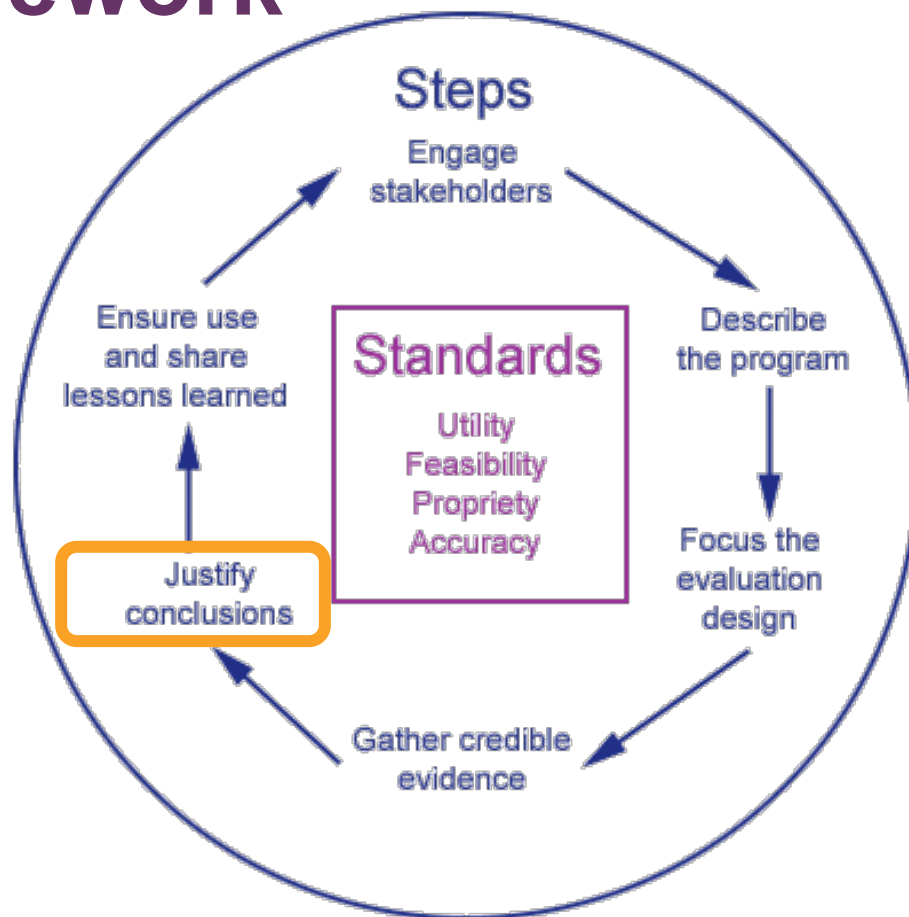
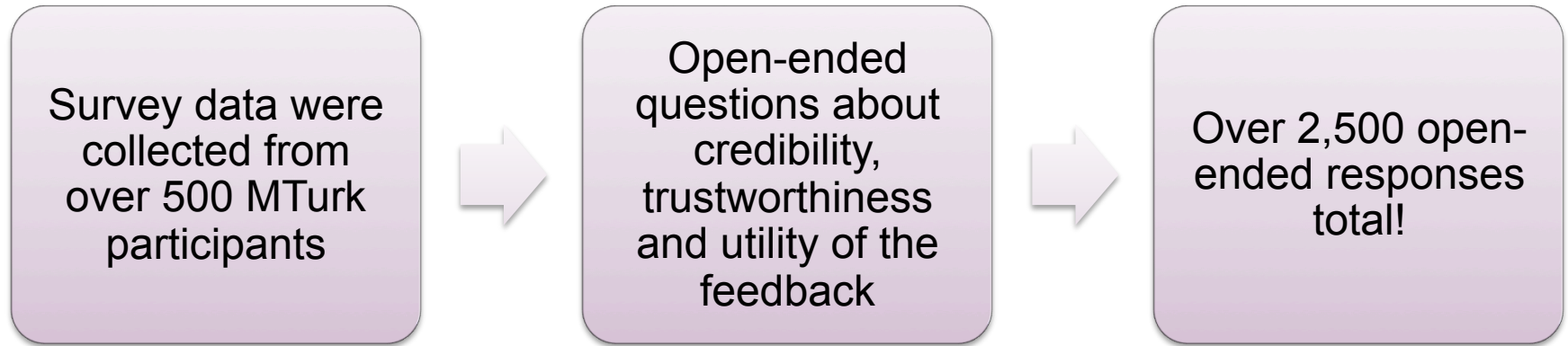

Engaging the crowd in analysis of qualitative survey responses

Cristina Whyte
Miriam Jacobson
Claremont Graduate University

CDC Framework



Researched best methods for delivering positive and negative feedback



**What do you do when you have
so many open-ended
responses?**

**What if we utilize the crowd to
code survey responses?**

Advantages of Utilizing the Crowd to Code



Saves Time



Cuts Costs



**Includes Alternative
Perspectives in Qualitative
Analysis**

Research Questions

1. What is the validity of coding qualitative survey responses in MTurk?
2. What is the feasibility of coding qualitative survey responses in MTurk?

Considerations for Task Development

Criteria for recruitment?

Amount to pay participants?

Revise the codes for the task?

How many codes?

How many responses should each participant code?

Include a practice portion?



Category	Definition	Example
Use of results	<p>Participants describe how they will use or not use the results to change their personal work ethic. For example, they might say that they will change how hard they work, how many hours they work, their attitude towards their work, how they handle obstacles, how they approach challenging situations, or how their work ethic will change by using the results.</p>	<p>“I will use those results to increase my work ethic. Since I was slightly above average, I want to continue working harder so I can be significantly above average.”</p> <p>“There's always room for improvement. I will use these results to motivate me to work harder.”</p>
Quality of personal work ethic	<p>Participants describe the quality of their current work ethic, but do not refer to the results or indicate whether they will use the results. Participants may describe how hard they work, how many hours they work, how often they work, the quality of their work, how coworkers or supervisors describe their work ethic, or how they handle obstacles.</p>	<p>“I’m motivated and goal driven.”</p> <p>“I tend to be lazy most days but feel good when I can work.”</p>
Agreement with results	<p>Participants describe the extent to which the results are consistent with their own view of their work ethic. Participants may agree or disagree with the results or comment on the extent to which the results are correct.</p>	<p>“The results are accurate based on my self-analysis.”</p> <p>“I believe I am higher in those traits than presented, but that they are not entirely untrue.”</p>

Practice

Please select the category that best fits each quote. Please review the description for each category below before starting the task.

	Use of results	Quality of personal work ethic	Agreement with results	Trustworthiness of results or survey	Other
I feel that I am very resilient, goal oriented, and a highly motivated person.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would describe myself as having a low work ethic, which is what was described in the results.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't plan on using the results for anything.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It's not a real indicator because its a survey.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
I think it can help people.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't have any reason for the results to lie to me so I believe them.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

The crowd and researchers coded 60 quotes obtained from two open-ended survey questions.

251 MTurk Workers

```
graph LR; A[251 MTurk Workers] --- B[Block of 20 Quotes (N=86)]; A --- C[Block of 20 Quotes (N=84)]; A --- D[Block of 20 Quotes (N=81)];
```

Block of 20 Quotes (N=86)

Block of 20 Quotes (N=84)

Block of 20 Quotes (N=81)

“I have no idea what I would use them for”

Code	# Mturkers	Researcher Code
Use of Results	64 ✓	✓
Quality of Work Ethic	5	
Agreement with Results	3	
Trustworthiness of Results/Survey	2	
Other	10	

Note: The quote was in response to the survey question: *Do you plan on using the results from the work ethic survey? Please explain.*

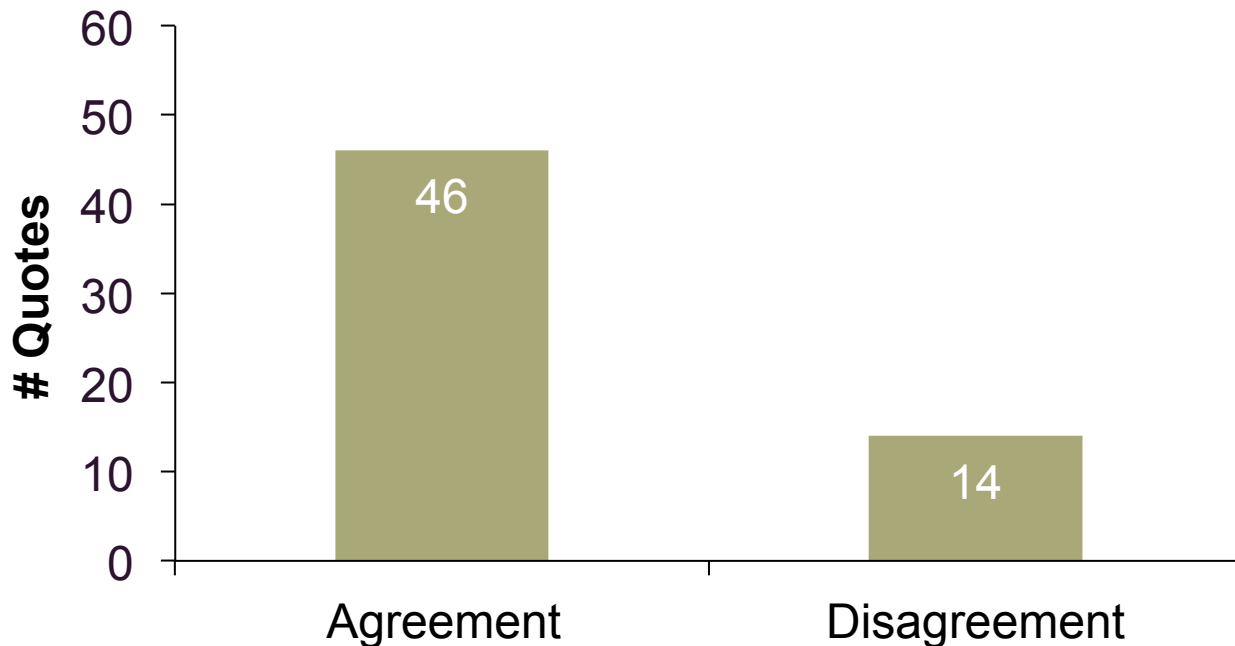
“Don’t think they are accurate”

Code	# Mturkers	Researcher Code
Use of Results	0	
Quality of Work Ethic	5	
Agreement with Results	35	✓
Trustworthiness of Results/Survey	38 ✓	
Other	6	

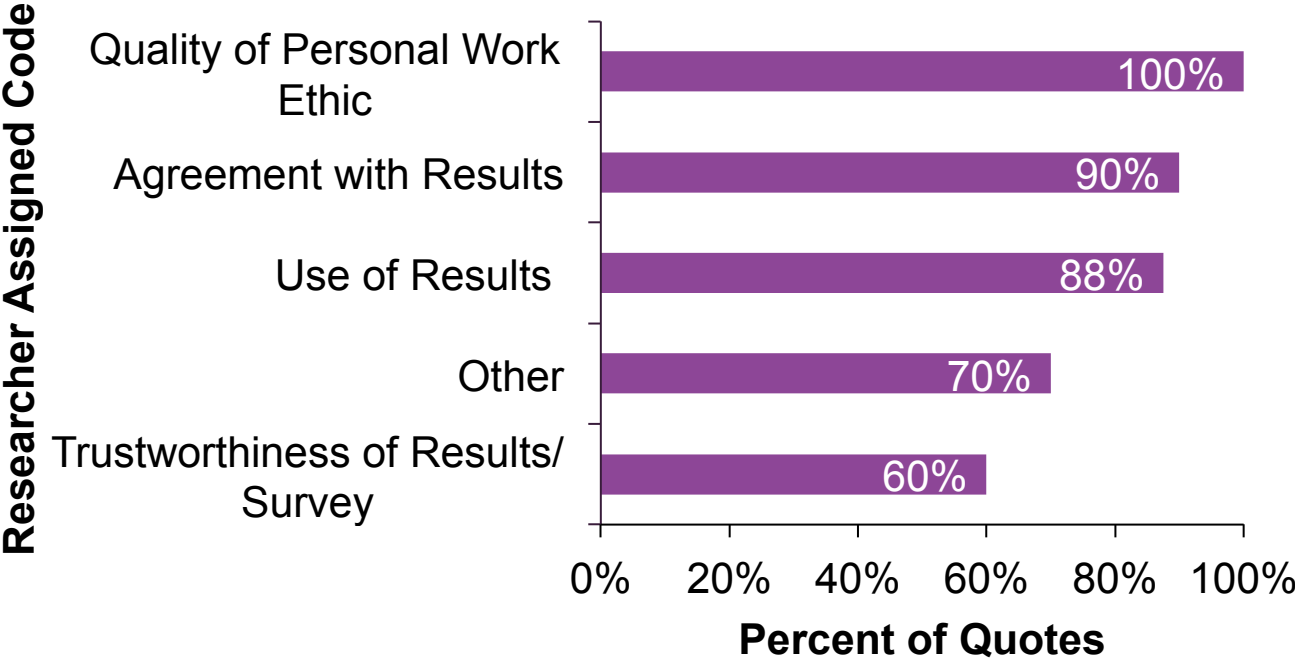
Note: The quote was in response to the survey question: *Do you plan on using the results from the work ethic survey? Please explain.*

Results: Agreement between the Crowd and Researchers

76.7% of quotes (46 out of 60) were coded consistently between the researchers and the crowd.



Percent of Quotes with Researcher-Crowd Agreement



Note: There were only 4 quotes coded as quality of personal work ethic

What is the validity and feasibility of crowdsourcing qualitative coding of survey responses?

Validity

- The crowd can be consistent with researchers.
- The crowd can provide feedback on coding schemes.

Feasibility

- 80+coders/20 quotes is too costly ($\approx \$115$)
- Time intensive to set-up

Thank you!

Cristina Whyte

cristina.tangonan@cgu.edu

Miriam Jacobson

miriam.jacobson@cgu.edu