

On Using Cognitive Computing and Machine Learning Tools to Improve Call Center Quality Control

FedCASIC 2019

Lew Berman, MS, PhD

John Boyle, PhD

Don Allen

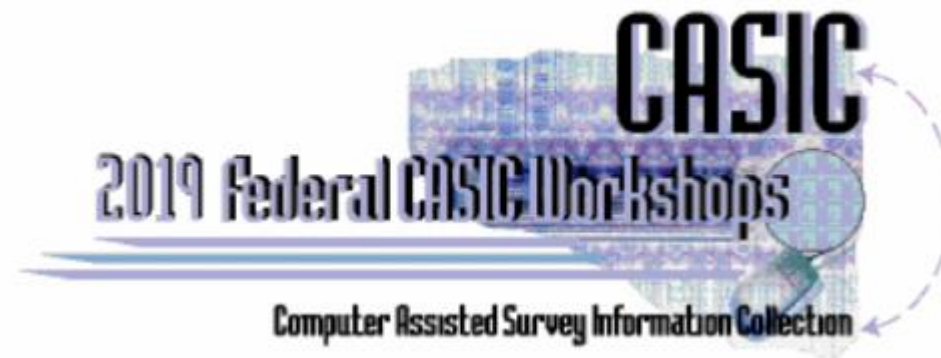
Josh Duell

Matt Jans, PhD

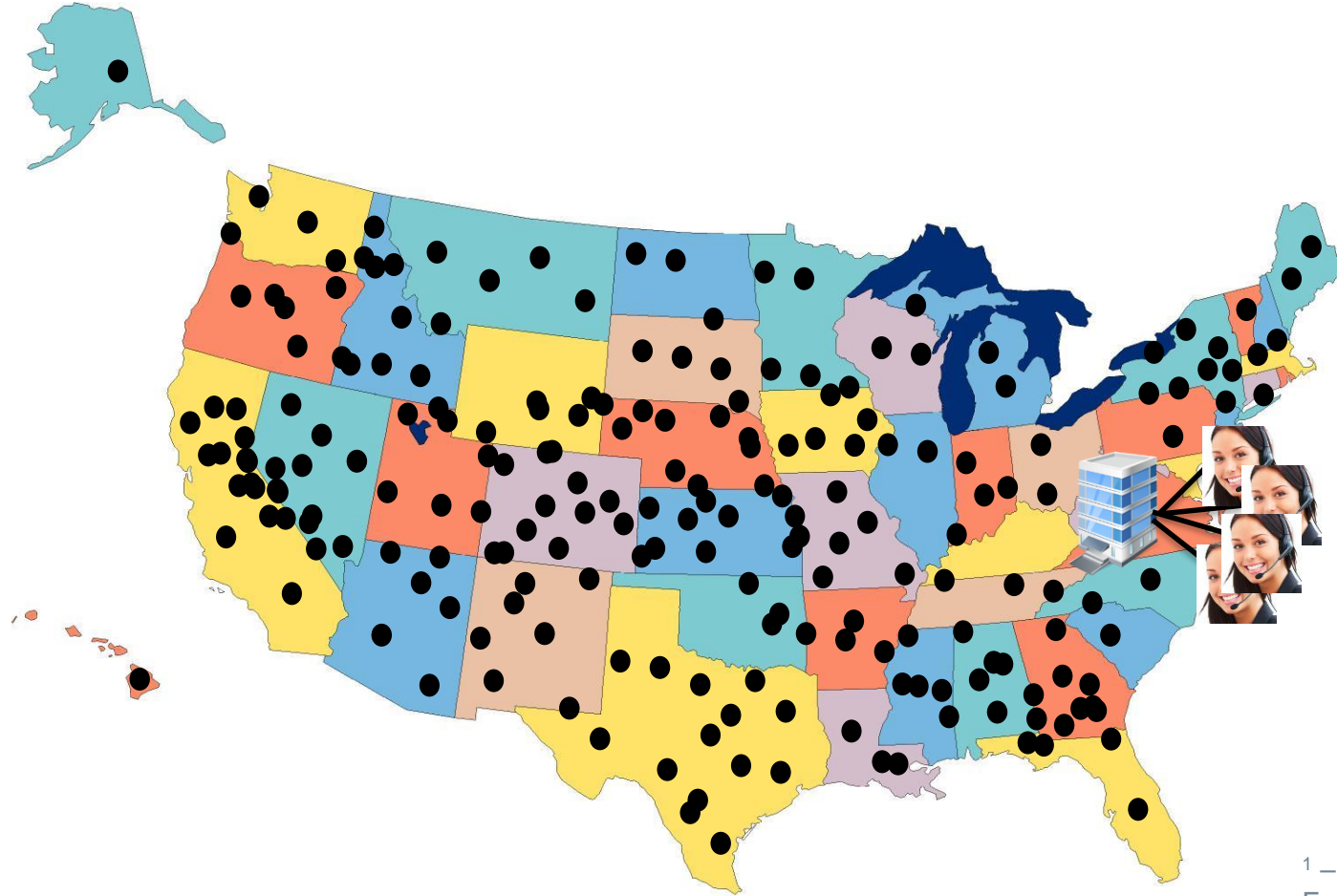
Ronaldo Iachan, PhD

Josiah McCoy

April 17, 2019



Use Case: Nationwide Telephone Survey



- Telephone-based survey
- Complex questions taking 20 minutes over the phone
- Topics may be sensitive
- Sample may include 1k – 10k respondents
- Quality review on 5-10% of sample¹

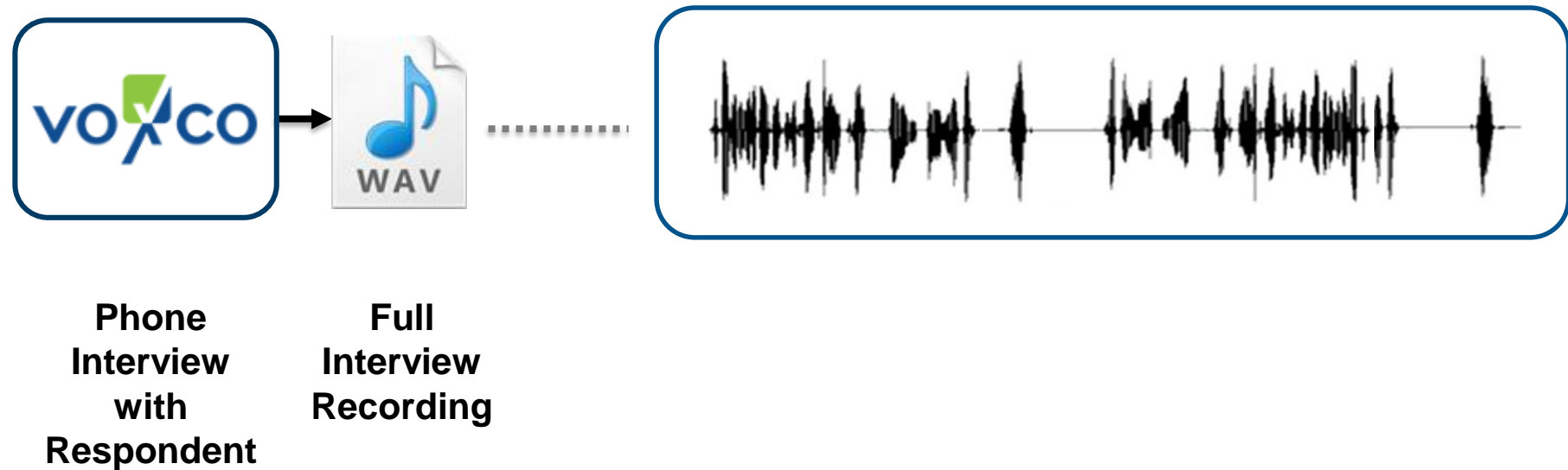
¹ – Roundtable: What Makes a Good Interviewer? Metrics and Methods for Ensuring Data Quality. FedCASIC 2018, Panel Moderated by Matt Jans, April 17-18, 2018.



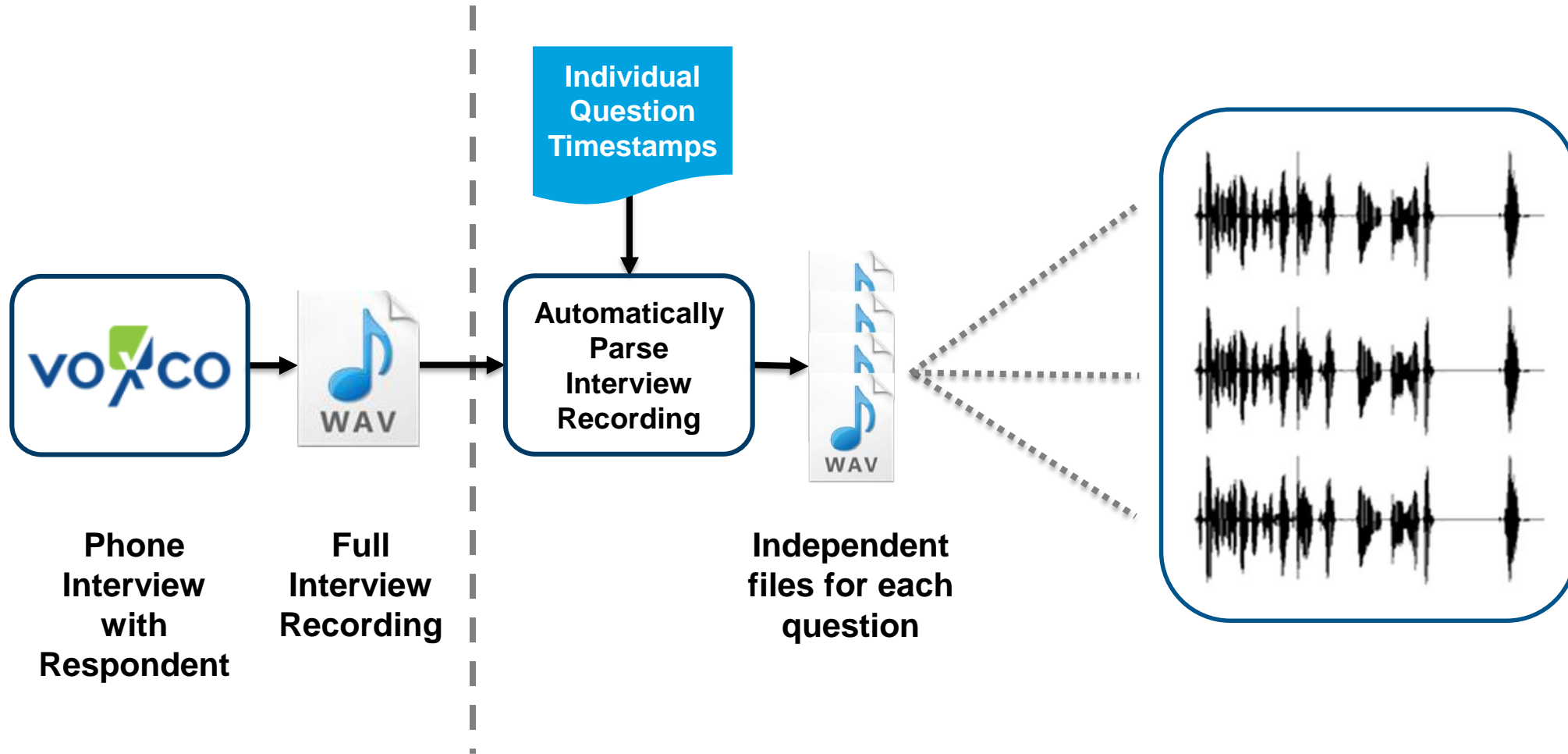
Problem: Which elements of Call Center QC can be automated and how?

- We can analyze audio recordings with Machine Learning and Speech Recognition software
- **Two questions:**
 - Can the software accurately identify which question is being asked?
 - Can the software accurately identify what the response was?
- **If we can do these two things, we can develop an automated process to check the accuracy of data recording at call centers**

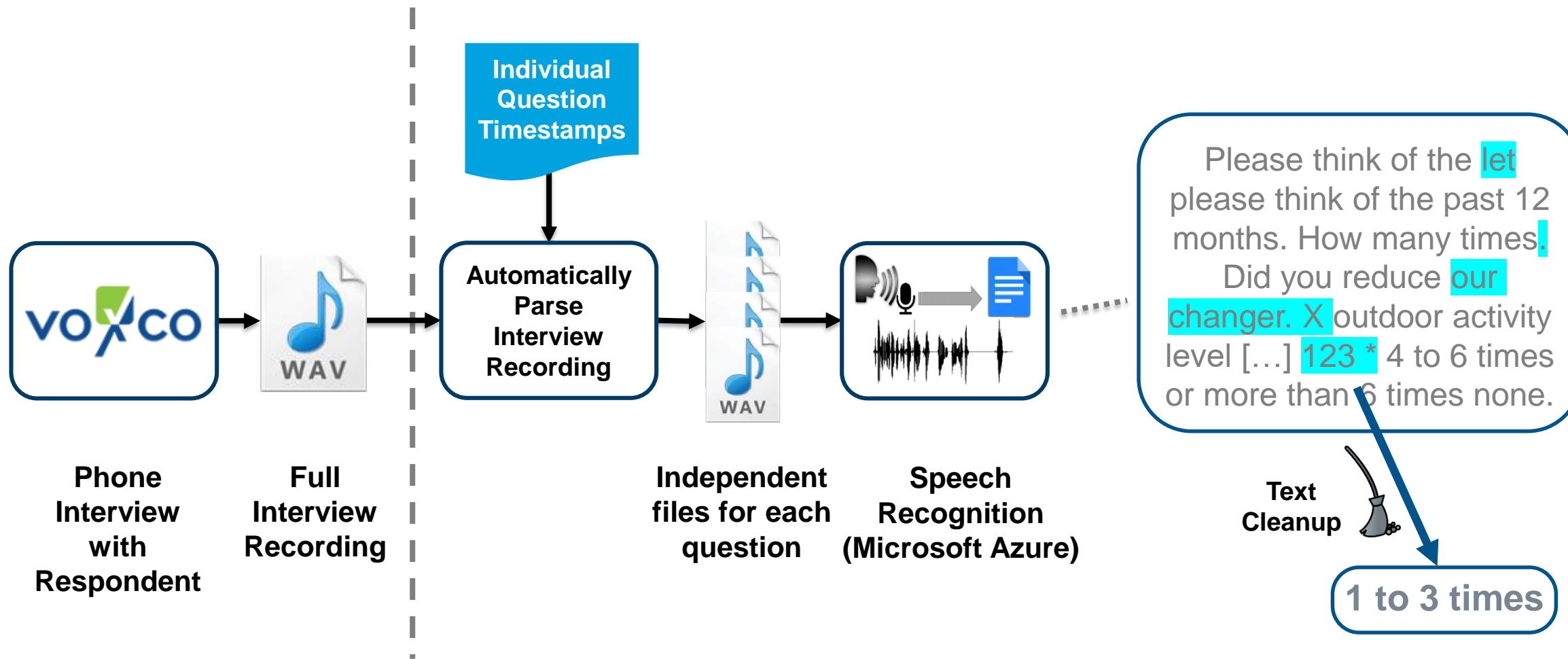
Overview of Speech Recognition and Language Understanding Process



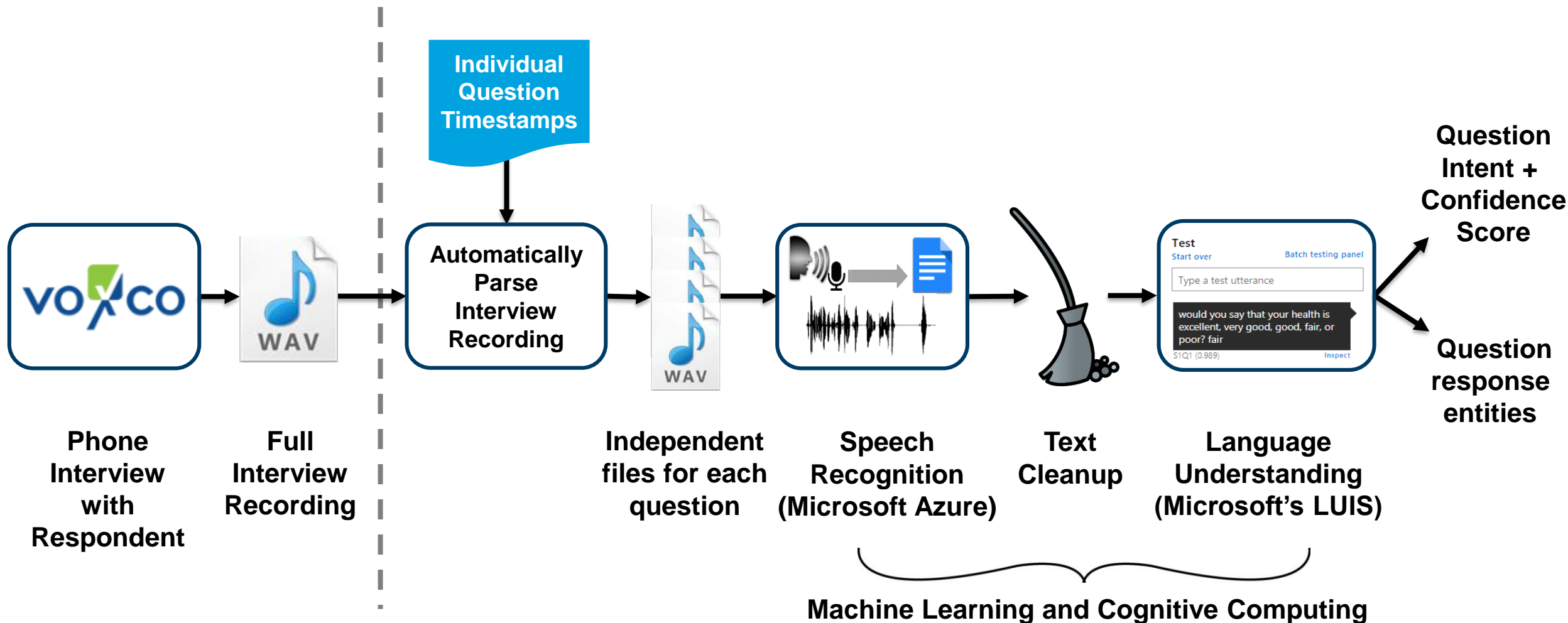
Overview of Speech Recognition and Language Understanding Process



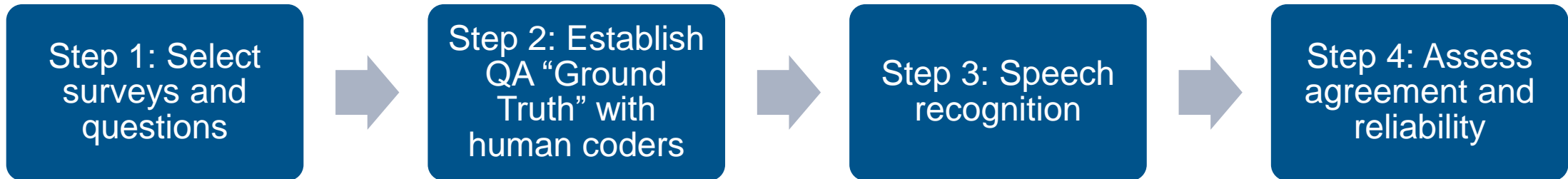
Overview of Speech Recognition and Language Understanding Process



Overview of Speech Recognition and Language Understanding Process



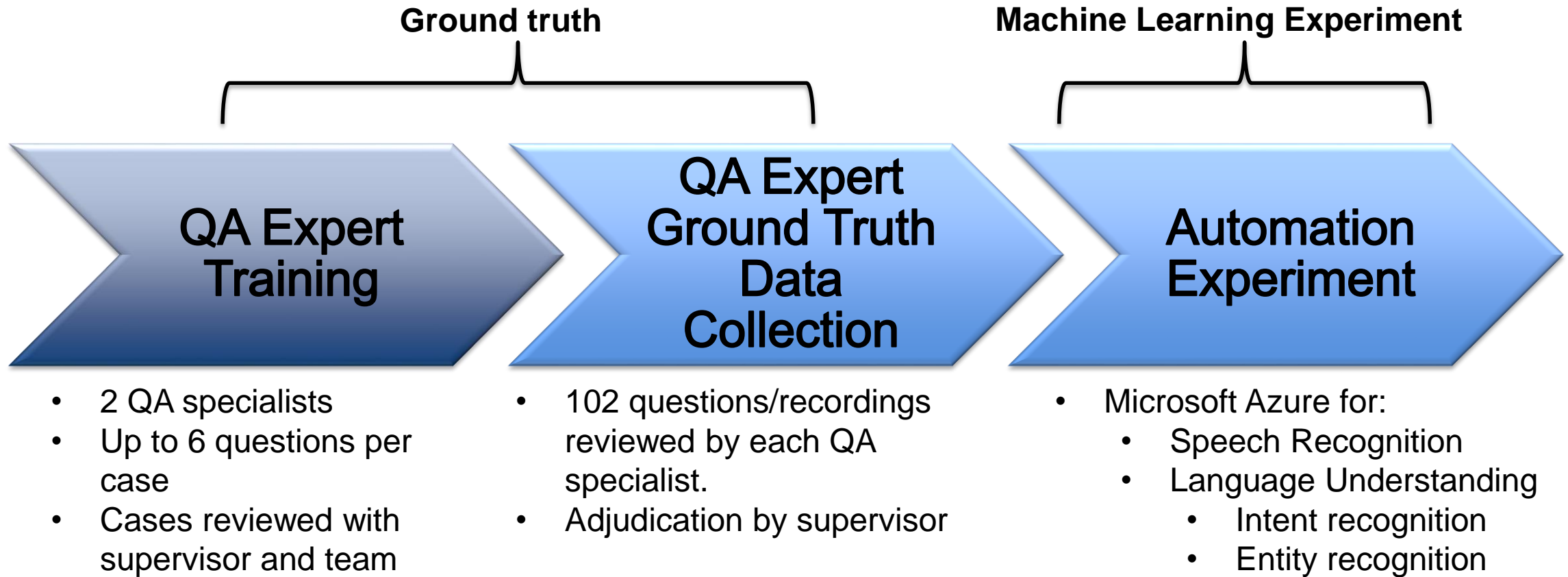
Methods Overview



Questions and Rationale for Automation

Popular question	Common response options	Question complexity	Response Complexity (e.g., many categories, likely probing)	Has definition or defined time frame
General Health: <i>Would you say that in general your health is excellent, very good, good, fair or poor? (S1Q1)</i>				
Y	Y (Vague quantifier)	Simple	Simple	N
Employment: <i>Are you currently employed for wages, self-employed, out of work for 1 year or more, out of work for less than 1 year, ...<u>or</u> unable to work? (S8Q15)</i>				
Y	N (Survey-specific)	Simple	Complex	N
Ever Smoker: <i>Have you smoked at least 100 cigarettes in your entire life? (S9q1)</i>				
Y	Y (Yes/No)	Simple	Complex	N
Seat Belts: <i>How often do you use seat belts when you drive or ride in a car? Would you say always, nearly always, sometimes, seldom, never? (S13Q1)</i>				
Y	Y (Vague quantifier)	Simple	Simple	N
Outdoor Exercise Change: <i>Please think of the past 12 months. How many times did you reduce or change your outdoor activity level based on the air quality index or air quality alerts? For example... (CT13_2)</i>				
N	Y (Numeric categories)	Complex	Simple (I) Complex (R)	Y

Experimental Steps





Reliability Results

(Adjudicated Ground Truth v.
Autocoded)

S1Q1. Would you say that in general your health is...

		Microsoft Language Understanding (Autocoded)					Total
		Excellent	Very Good	Good	Fair	Poor	
Ground Truth (Adjudicated)	Excellent	4	0	0	0	0	4
	Very Good	0	10	0	0	0	10
	Good	0	0	4	0	0	4
	Fair	0	0	0	4	1	5
	Poor	0	0	0	0	0	0
	Total	4	10	4	4	1	23

S1Q1. Would you say that in general your health is:

- 1 Excellent
- 2 Very good
- 3 Good
- 4 Fair
- 5 Poor

Agreement	Kappa	Std. Error	Z	Prob > Z
95.65%	0.9390	0.1205	7.79	0.000

Reliability Across Questions

<u>Question)</u>	<u>Type</u>	<u>Agreement</u>	<u>Kappa</u>
General Health (s1q1)	Simple question, Simple response	96%	0.94
Employment (s8q15)	Simple question, Complex response	61%	0.50
100 Cigarettes (s9q1)	Simple question, Complex response	70%	0.53
Seat belts (s13q1)	Simple question, Simple response	70%	0.0
Outdoor activity change (CT13_2)	Complex (and long question Simple response (I) Complex response (R)	20%	0.0



Summary

- **Successful first steps!**
- **High reliability for simple questions with simple response**
- **Reliability can be low if question is complex, even if responses seems simple (i.e., outdoor exercise change)**



Limitations and Challenges

- **Voxco (and other systems) typically record interview as one “.wav” file**
 - Manual splitting required to one recording per question
- **Required some adjustment to set for file type and sampling rate**
- **Future work required to differentiate**
 - Signal & noise
 - Interviewer, respondent, dog, baby, ...



Discussion – Human Impacts

▪ Interviewers

- Provide timely feedback for new interviewers
- Follow-up with objective information for interviewers with poor performance
- Should not be used in a punitive manner, but as a tool for corrective action

▪ Management

- Understand impact of coaching and training
- Produce daily, weekly, monthly trends reports for supervisors, interviewers, and clients

▪ Sponsor perspective

- Increase level of monitoring
- Focus on key interview questions for 100% monitoring

Thank you!

- Lew Berman (lew.berman@icf.com)
- Josiah McCoy (josiah.mccoy@icf.com)
- Matt Jans (matt.jans@icf.com)

- **Acknowledgements**
 - Stephanie Florence, Quality Assurance Specialist
 - JaWanda Martin, Quality Assurance Specialist

s8q15: Employment

```
RESPONSE_num_S8Q15_r |
ec: S8Q15 response |
from Adjudication |
file (numeric, | s8q15 response from auto system (recoded for kappa)
recoded) | 1: Emp fo 2: Self-e 4: Out-of 7: Retire 8: Unable | Total
```

	1: Emp fo	2: Self-e	4: Out-of	7: Retire	8: Unable	Total
1: Emp for wages	4	1	0	0	3	8
2: Self-emp	0	1	0	0	1	2
4: Out-of-work < 1yr	0	0	0	0	2	2
7: Retired	0	0	0	7	2	9
8: Unable to work	0	0	0	0	2	2
Total	4	2	0	7	10	23

Agreement	Expected Kappa	Std. Err.	Z	Prob>Z
60.87%	22.50%	0.4951	0.1007	4.92 0.0000

S8q15: Are you currently...?

INTERVIEWER NOTE: If more than one, say "Select the category which best describes you".

- PLEASE READ:**
- 1 Employed for wages
 - 2 Self-employed
 - 3 Out of work for 1 year or more
 - 4 Out of work for less than 1 year
 - 5 A Homemaker
 - 6 A Student
 - 7 Retired
 - Or
 - 8 Unable to work
- DO NOT READ:**
- 9 Refused

S8Q15. Are you currently...

		Microsoft Language Understanding Response							Total	
		Employed for Wages	Self-Employed	Out of Work 1+ yr	Out of Work <1 yr	Homemaker	Student	Retired		Unable to Work
Ground Truth	Employed for Wages	4	1	0	0	0	0	0	3	8
	Self-Employed	0	1	0	0	0	0	0	1	2
	Out of Work 1+ yr	0	0	0	0	0	0	0	0	0
	Out of Work <1 yr	0	0	0	0	0	0	0	2	2
	Homemaker	0	0	0	0	0	0	0	0	0
	Student	0	0	0	0	0	0	0	0	0
	Retired	0	0	0	0	0	0	7	2	9
	Unable to Work	0	0	0	0	0	0	0	2	2
	Total	4	2	0	0	0	0	7	10	23

S8q15: Are you currently...?

- 1 Employed for wages
- 2 Self-employed
- 3 Out of work for 1 year or more
- 4 Out of work for less than 1 year
- 5 A Homemaker
- 6 A Student
- 7 Retired, Or
- 8 Unable to work

DO NOT READ:

- 9 Refused

Agreement	Expected Agreement	Kappa	Std. Error	Z	Prob > Z
-----------	--------------------	-------	------------	---	----------

60.87% 22.50% 0.4951 0.1007 4.92 0.000

s9q1: 100 Cigarettes

s9q1 response from auto system (recoded for kappa)					
RESPONSE_num_	S9Q1_4cat	1: Yes	2: No	97: Uncod	Total
1: Yes		8	0	3	11
2: No		0	8	4	12
97: Uncodable		0	0	0	0
Total		8	8	7	23
Agreement	Expected Kappa	Std. Err.	Z	Prob>Z	
69.57%	34.78%	0.5333	0.1332	4.00	0.0000

S9q1: Have you smoked at least 100 cigarettes in your entire life?

INTERVIEWER NOTE: For cigarettes, do not include: electronic cigarettes (e-cigarettes, njoy, bluetip), herbal cigarettes, cigars, cigarillos, little cigars, pipes, bidis, kreteks, water pipes (hookahs) or marijuana.

INTERVIEWER NOTE: 5 packs = 100 cigarettes

- 1 Yes
- 2 No [Go to S9Q5]
- 7 Don't know / Not sure [Go to S9Q5]
- 9 Refused [Go to S9Q5]

s13q1: Seat belts

RESPONSE_num_S13	Q1_rec: S13Q1	response from	Adjudication	file (numeric, recoded)	s13q1 response from auto system (recoded for kappa)	Total
					1: Always 2: Nearly 3: Someti 5: Never 97: Uncod	
1: Always					16 1 1 4 1	23
2: Nearly Always					0 0 0 0 0	0
3: Sometimes					0 0 0 0 0	0
5: Never					0 0 0 0 0	0
97: Uncodable					0 0 0 0 0	0
Total					16 1 1 4 1	23
Agreement	Expected Kappa	Std. Err.	Z	Prob>Z		
69.57%	69.57%	0.0000	0.0000	.	.	

s13q1: How often do you use seat belts when you drive or ride in a car? Would you say—

PLEASE READ:

- 1 Always
- 2 Nearly always
- 3 Sometimes
- 4 Seldom
- 5 Never

DO NOT READ:

- 7 Don't know / Not sure
- 8 Never drive or ride in a car
- 9 Refused

CT13_2: Outdoor activity change

RESPONSE_num_	CT13_2_rec:	CT13_2	response from	Adjudication	file	CT13_2 response from auto system (recoded for kappa)	recode	1: None	2: 1-3 ti	3: 4-6 ti	4: >6 tim	Total
1: None								1	1	1	0	3
2: 1-3 times								0	0	0	0	0
3: 4-6 times								0	1	0	0	1
4: >6 times								0	0	1	0	1
Total								1	2	2	0	5
Agreement	Expected Kappa	Std. Err.	Z	Prob>Z								
20.00%	20.00%	0.0000	0.1732	0.00	0.5000							

CT13_2: Please think of the past 12 months. How many times did you reduce or change your outdoor activity level based on the air quality index or air quality alerts? For example avoiding outdoor exercise or strenuous outdoor activity. Please do not include times when you may have heard or read about high pollen counts. Would you say:

PLEASE READ:

- 1. None
- 2. 1 to 3 times
- 3. 4 to 6 times or
- 4. More than 6 times

DO NOT READ:

- 7. Don't know/Not sure
- 9. Refused