



IVIP

**The effect of voice
sample duration and
lineup size on voice
identification
performance**

Harriet M. J. Smith, Nikolas Pautz, Kirsty E.
McDougall, Katrin Mueller-Johnson, Alice
Paver, & Francis Nolan





Voice identification evidence

ID evidence can be determinative

In some crimes visual information is not available

Voice identification is admissible evidence in jurisdictions worldwide

What are the main issues?

High false alarm rates (Kerstholt et al., 2004; Stevenage et al., 2012, 2013)

Juries find voice identification evidence extremely persuasive (Van Wallendael et al., 1994)

Unfamiliar voice identification is under-researched, especially when it comes to system variables

Improving voice identification procedures (IVIP)



ESRC-funded project (ES/S015965/1)

Multi-disciplinary approach (psychology, forensic phonetics, linguistics, criminology & law)

4 strands:

- Strand 1: What are the current voice identification procedures?
- Strand 2: What are the factors that affect the reliability of voice identification?
- Strand 3: How do we improve the reliability of voice identification?
- Strand 4: How accurate are current voice identification procedures used by criminal justice practitioners in response to voice identification requests?



Home Office circular 057/2003: 'Advice on the use of voice identification procedures'



1. Representative sample of the suspect speaking naturally.
2. Voice samples should be 1 minute long
3. Voice parade should consist of 9 voices
4. Witness must be instructed that the voice of the suspect may or may not be present
5. The witness must listen to each tape at least once before making a selection
6. The witness must be allowed to listen to the samples as many times as they wish

Experiment 1



Can sample durations be reduced without a performance cost?

Practical considerations - time consuming for the police

Voice samples should
be 1 minute long

People can extract basic identity information
from much shorter durations (Bestelmeyer et al., 2010;
McAleer et al., 2014)

Voice parade will take at least 9 minutes

Temporal ratio models of memory (Bjork and Whitten,
1974; Brown et al., 2007)

Speaker / Stimulus selection

In 'real' voice lineups recordings of the suspect and foil voices are taken from recordings of police interviews

Forensics-orientated speech databases

Dynamic Variability in Speech (**DYVIS**)

York Variation in Speech (**YORVIS**)

West Yorkshire Regional English Database (**WYRED**)

These databases include **recorded telephone calls** of a perpetrator discussing a crime and **mock police interviews**

Speaker	Distance
s010	0.00
s01	1.81
s013	1.94
s05	2.70
s07	2.87
s09	2.91
s014	3.19
s04	3.29
s011	3.30
s03	3.34
SUM-1-10	25.37
s08	3.37
s06	3.41
s02	3.89
s012	3.93
s015	4.08

Experiment 1 $N = 271$ (135 female)



IVs

Lineup sample: 15s, 30s or 60s

Perpetrator: present or absent

DVs

Accuracy

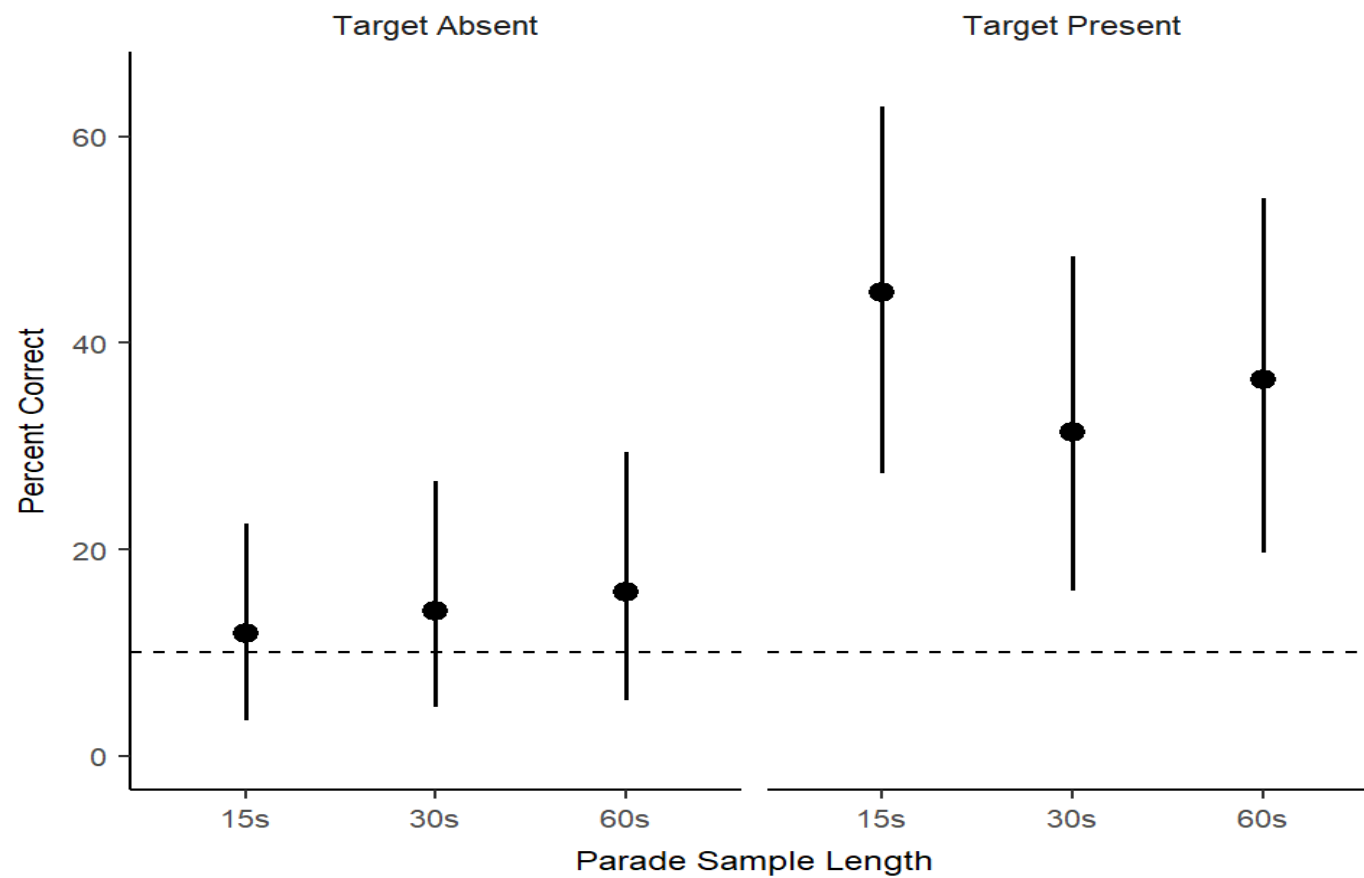
Self-rated confidence
(0-10)

Decision frequency

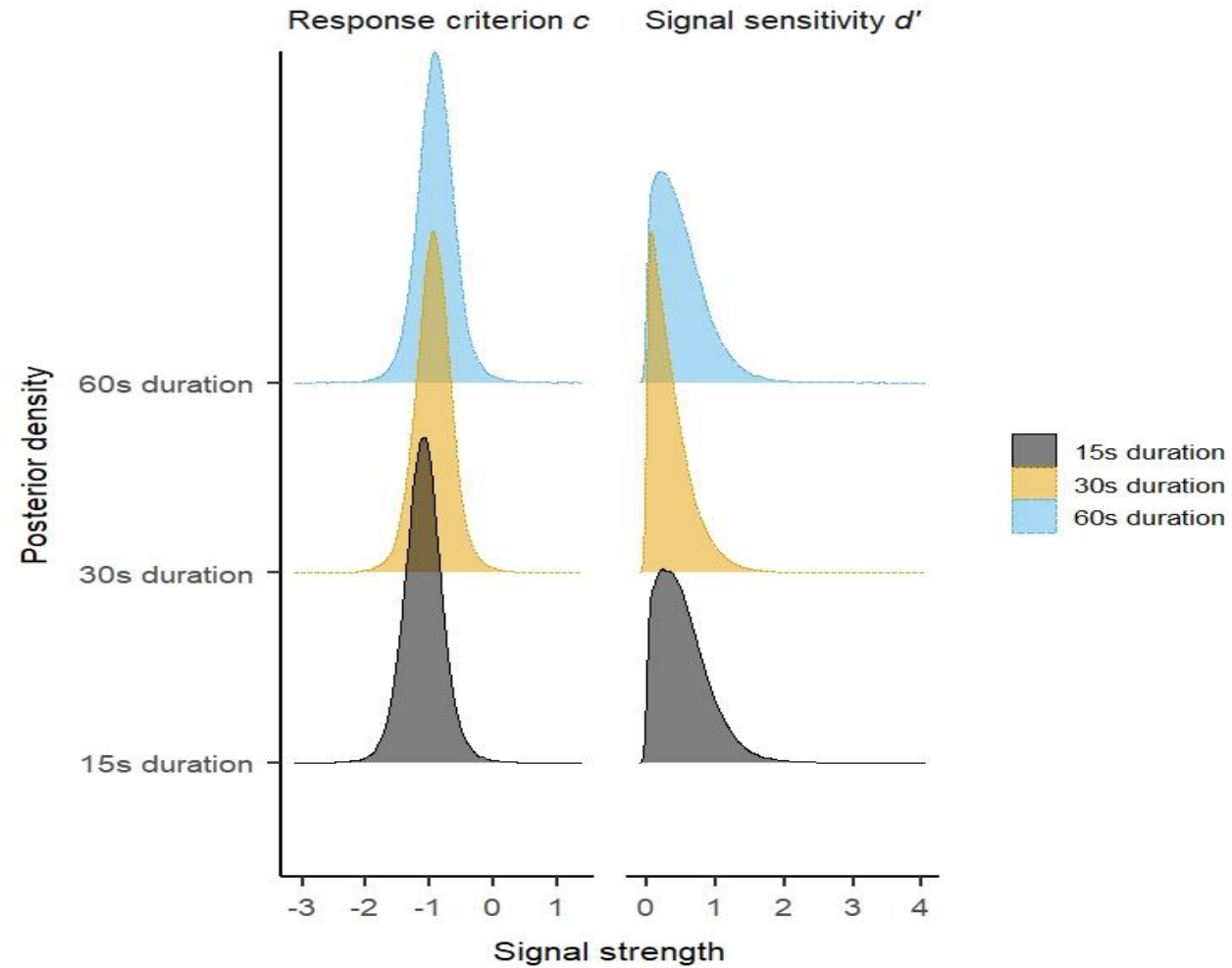


Sample Duration	Target Present			Target Absent	
	Hits	Foil	Reject	Foil	Reject
15 seconds	20 (45%)	21 (48%)	3 (7%)	42 (88%)	6 (13%)
30 seconds	14 (32%)	26 (59%)	4 (9%)	40 (85%)	7 (15%)
60 seconds	17 (37%)	27 (59%)	2 (4%)	35 (83%)	7 (17%)
Total	51 (38%)	74 (55%)	9 (7%)	117 (85%)	20 (15%)

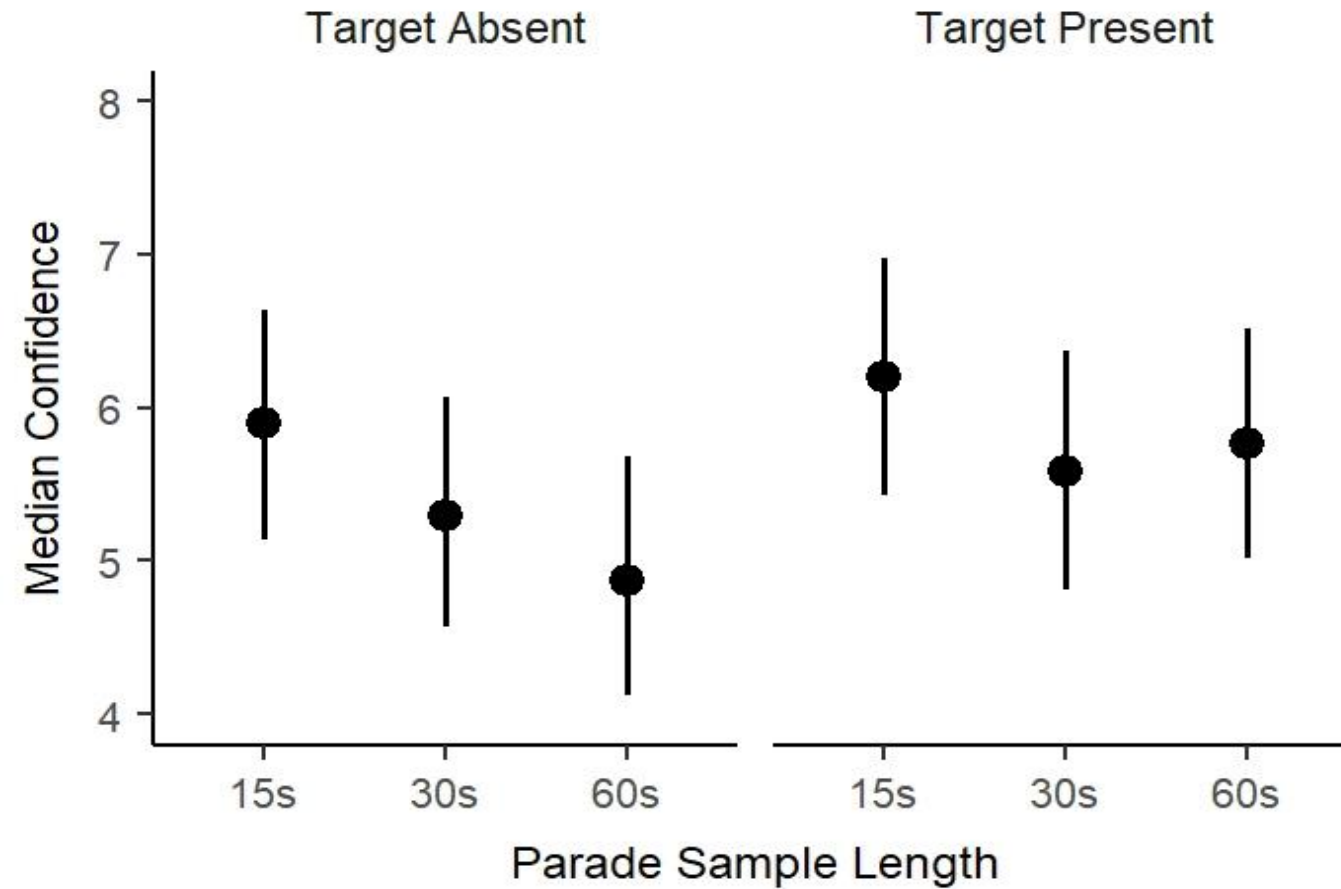
Accuracy



SDT analyses



Confidence



Experiment 1 conclusions



The results underline the importance of admitting voice identification with caution. Consistent with previous research, performance is low.

Our results highlight the importance of stimulus sampling. Many earwitness studies have been conducted using single targets.

No evidence to suggest that there is any benefit in using lineup samples of 60 s. These preliminary results suggest that the voice identification procedure currently recommended in England and Wales can be safely adapted by reducing the duration to 30 s or even 15 s.

Experiment 2



Can lineup size be reduced without a performance cost?

Voice parades should
consist of 9 voices

Practical considerations

Larger lineups offer more protection to
innocent suspect?

But erroneous auditory attention is more
likely to occur when the demand on
resources is high (Zimmerman, Moscovitch & Alain, 2016)

Experiment 2 $N=278$ (136 female)



IVs

Lineup sample: 15s, 30s or 60s

Perpetrator: present or absent

DVs

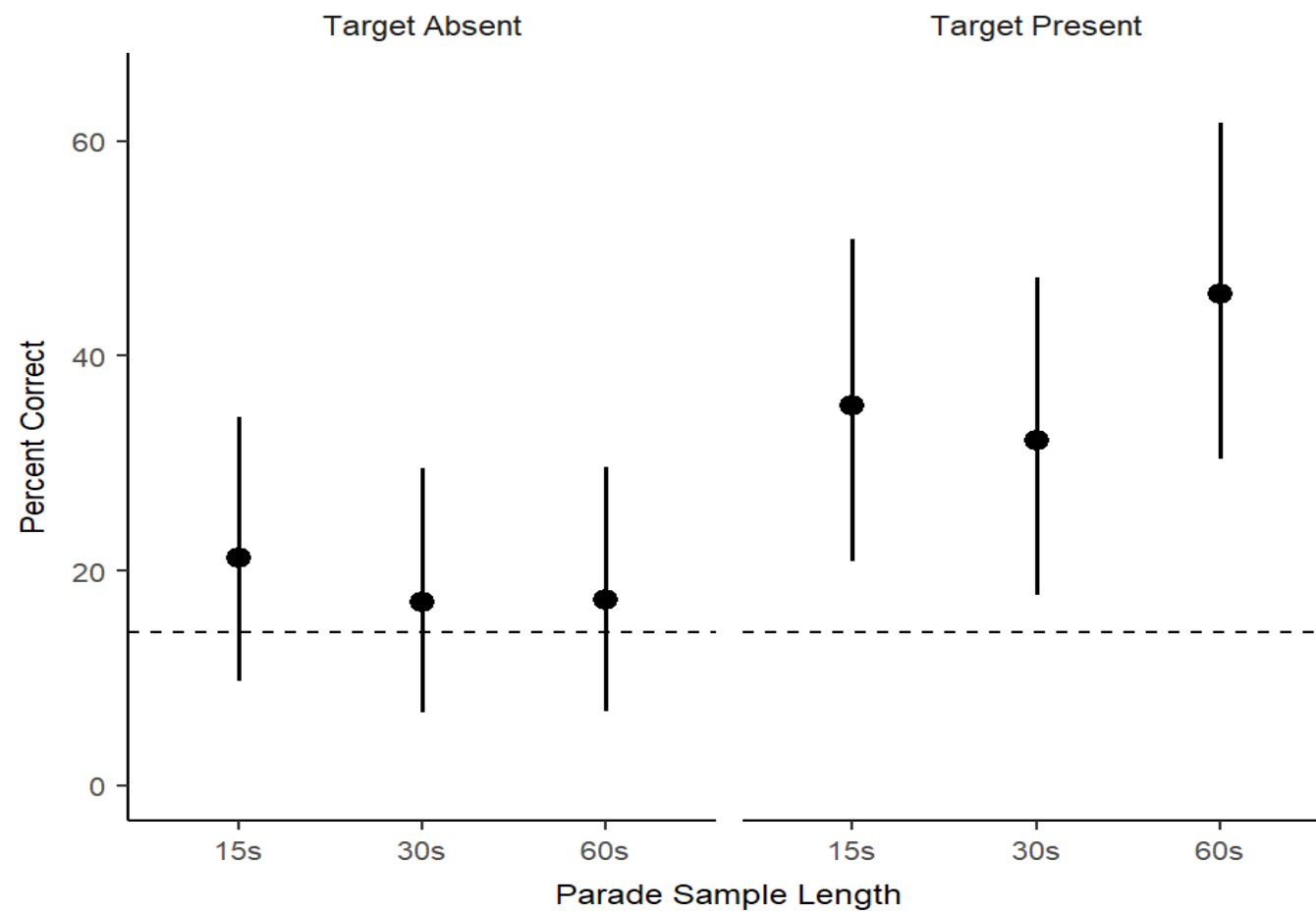
Accuracy

Self-rated confidence
(0-10)

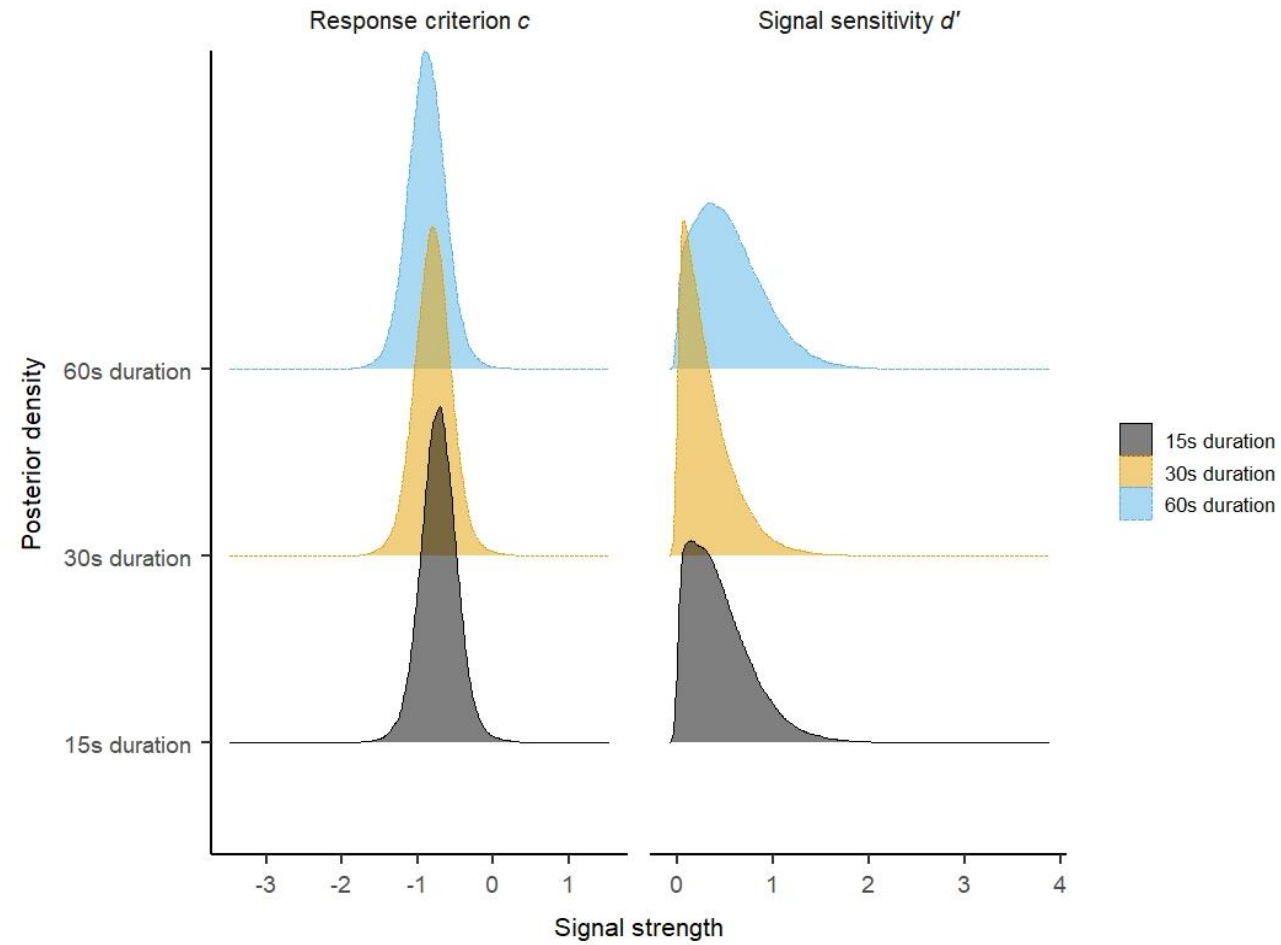
Decision frequency

Sample Duration	<u>Target Present</u>			<u>Target Absent</u>	
	Hit	Foil	Reject	Foil	Reject
15 seconds	16 (36%)	26 (58%)	3 (7%)	36 (78%)	10 (22%)
30 seconds	14 (33%)	22 (51%)	7 (16%)	37 (82%)	8 (18%)
60 seconds	21 (46%)	21 (46%)	4 (9%)	37 (82%)	8 (18%)
Total	51 (38%)	69 (51%)	14 (10%)	110 (81%)	26 (19%)

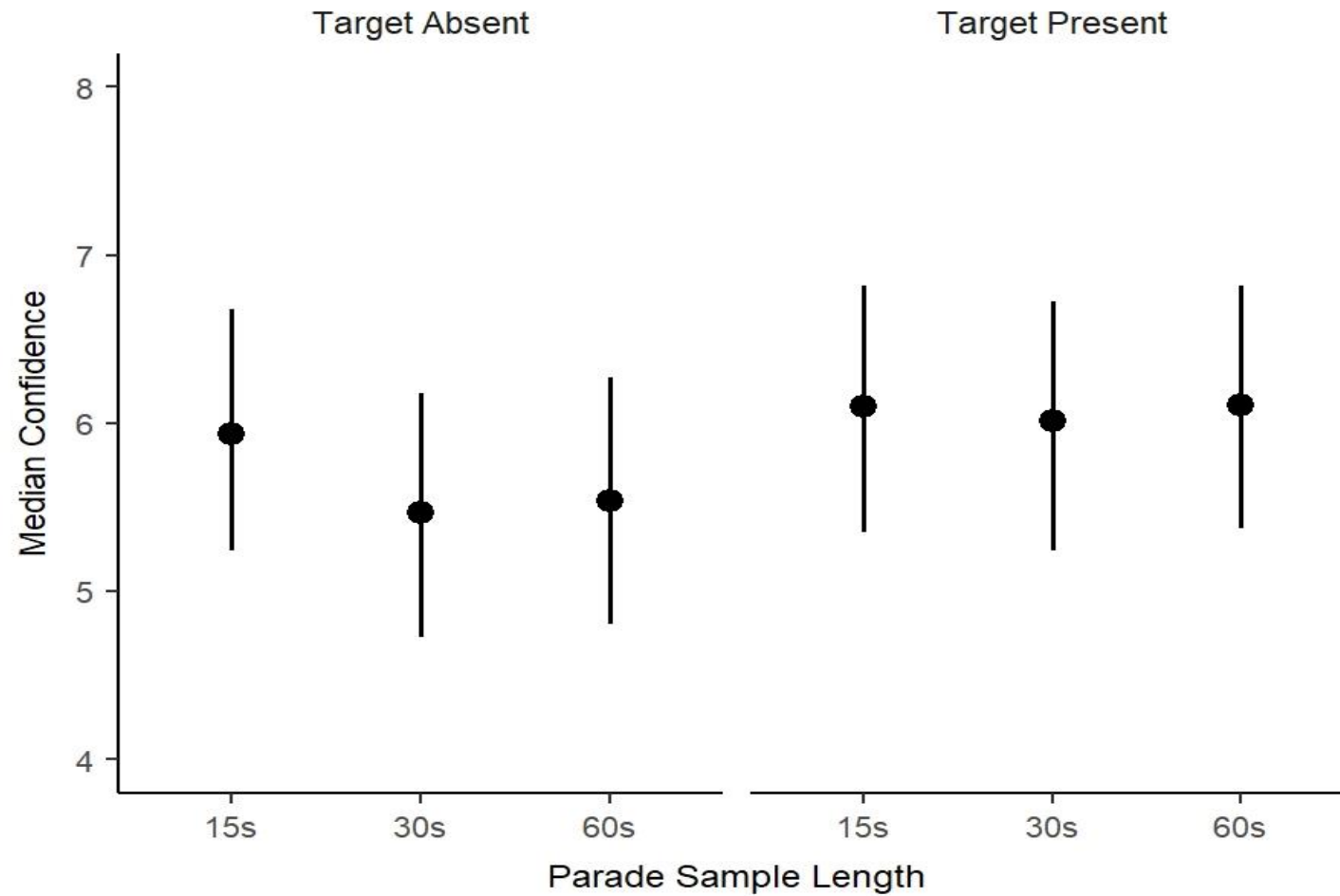
Accuracy



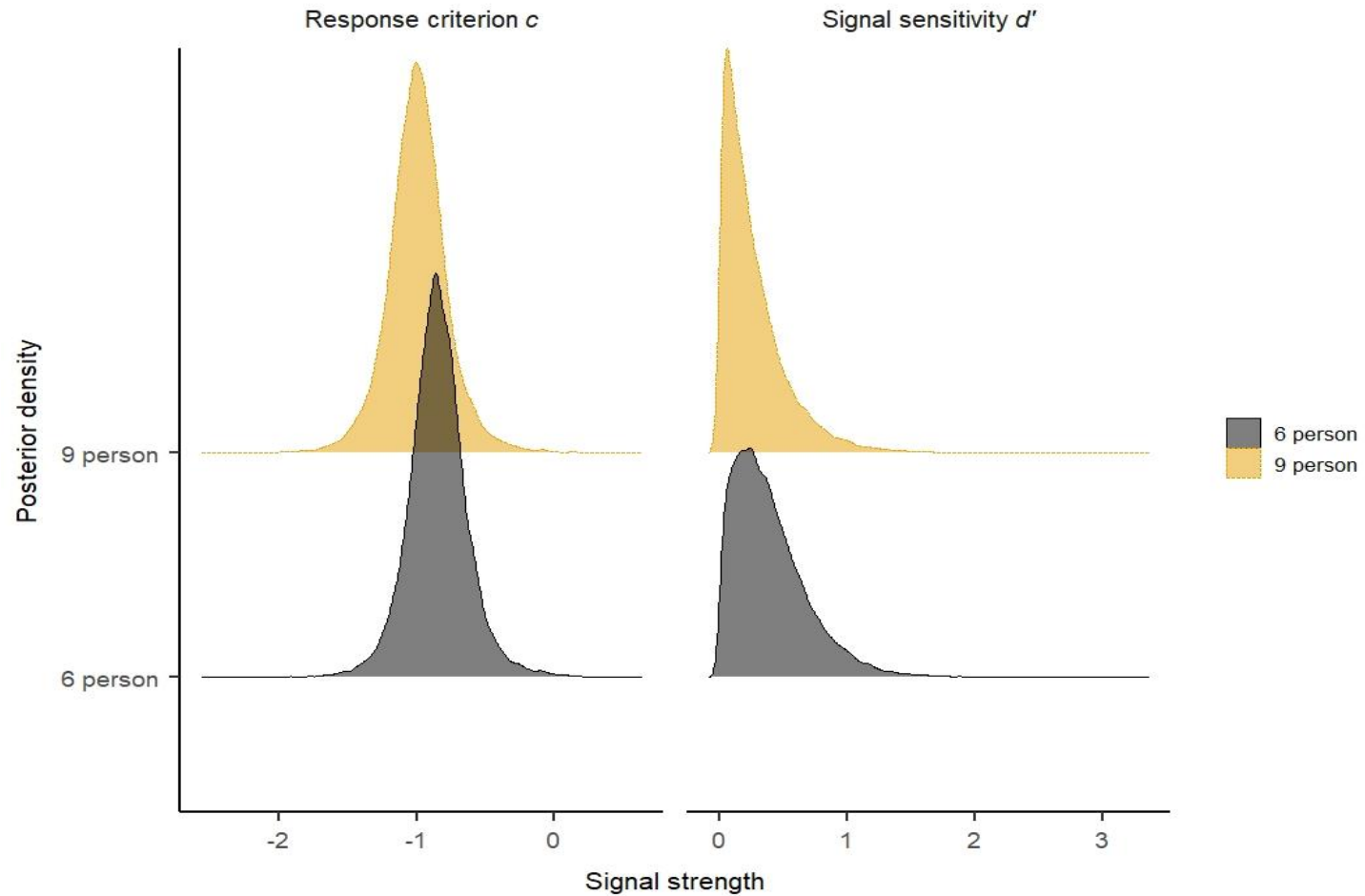
SDT analyses



Confidence



Experiment 1 and 2 comparison



Conclusions



Poor performance – but the task is not impossible if the target is present

Reduce sample duration? 

Reduce the number of foils? 



IVIP

Thank you for listening



@harrietsmith15



Harriet M. J. Smith



harriet.smith02@ntu.ac.uk

