KNOWING IF THEY KNOW:
A NOVEL BIAS-FREE METHOD FOR INCENTIVISING ACCURATE METACOGNITIVE REPORTS

Ryan Scott[1,2]
Bojana Ivic[1]
Zoltan Dienes[1,2]

[1]School of Psychology, University of Sussex
[2]Sackler Centre for Consciousness Science

METACOGNITION

Often operationalised using confidence

- Objective decision accuracy – demonstrates knowing
- Confidence-accuracy correlation – demonstrates metacognition

METACOGNITION

Often operationalised using confidence

- Objective decision accuracy – demonstrates knowing
- Confidence-accuracy correlation – demonstrates metacognition

Rationale for using confidence

- For both subliminal perception (Merikle, '07) and implicit leaning (Dienes, '08) 'awareness' is definitional of the phenomena we study
- If you endorse HOT theory then this is definitional of consciousness
- Confidence ratings directly evaluate awareness while better meeting the information criterion (Shanks & St John, '94)
- Demonstrable improvement over free report (Ziori & Dienes, 2006)

METACOGNITION

Often operationalised using confidence

- Objective decision accuracy – demonstrates knowing
- Confidence-accuracy correlation – demonstrates metacognition

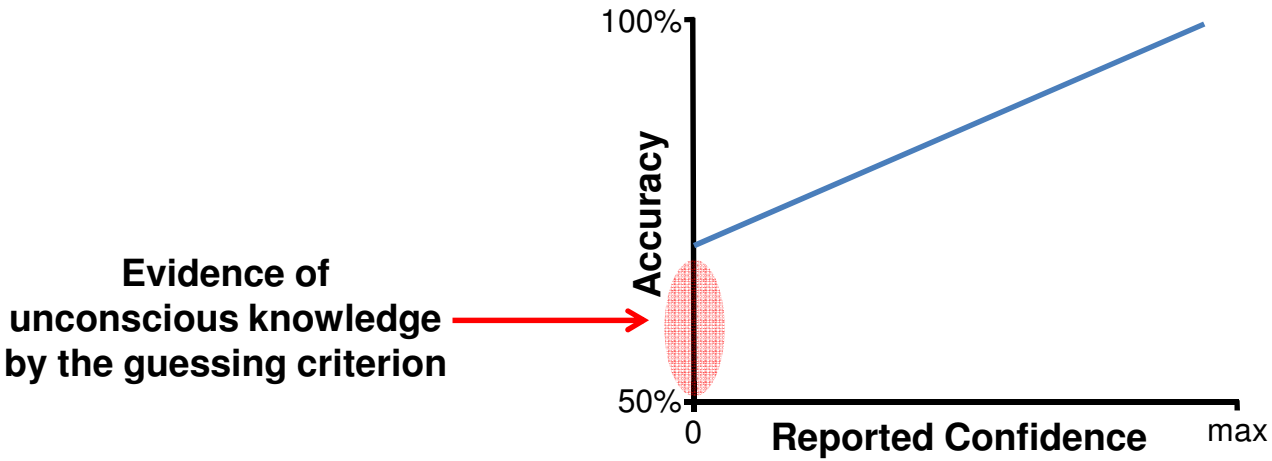
Rationale for using confidence

- For both subliminal perception (Merikle, '07) and implicit leaning (Dienes, '08) 'awareness' is definitional of the phenomena we study
- If you endorse HOT theory then this is definitional of consciousness
- Confidence ratings directly evaluate awareness while better meeting the information criterion (Shanks & St John, '94)
- Demonstrable improvement over free report (Ziori & Dienes, 2006)

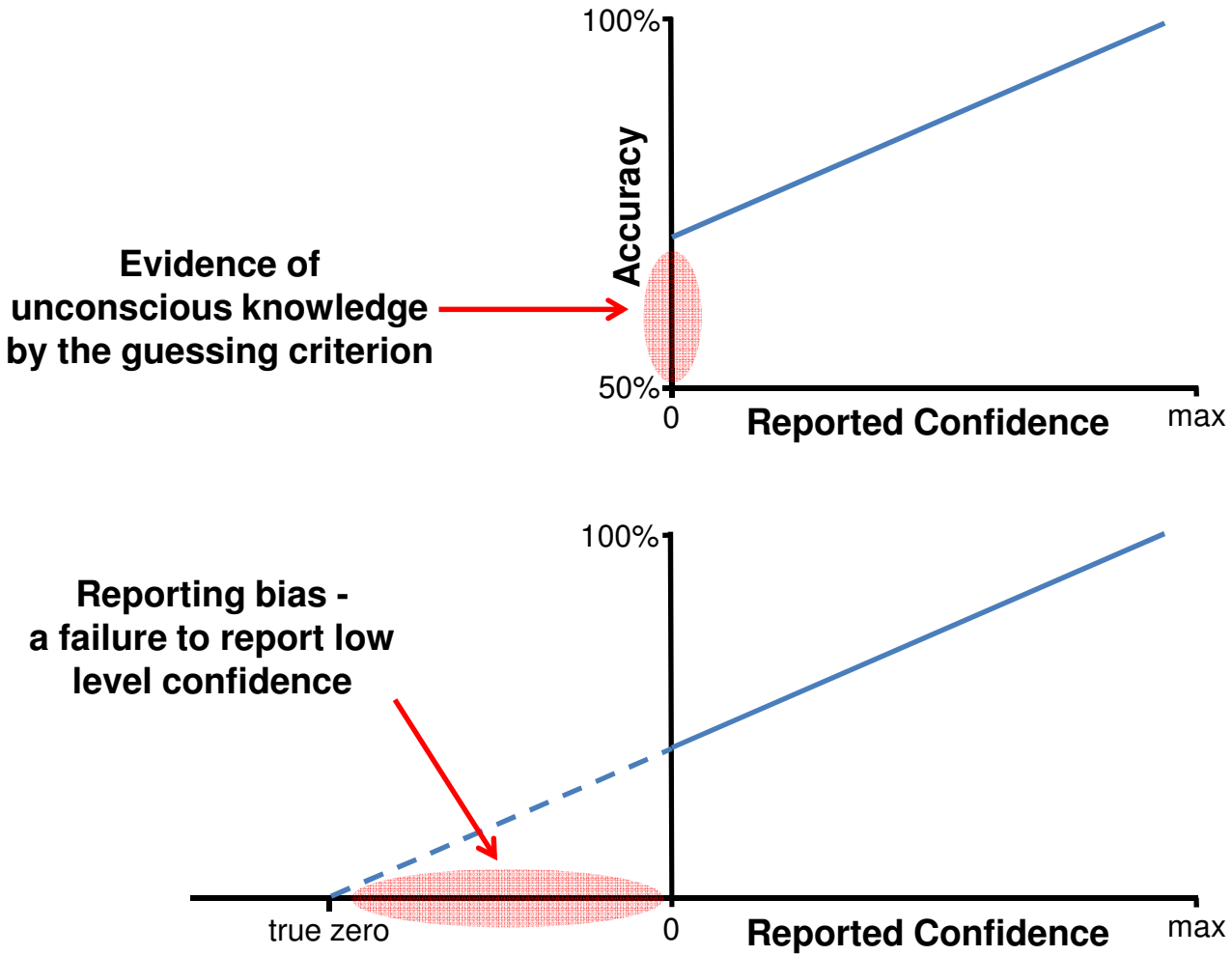
Potential limitations

- Bias – confidence itself may be under reported
- For some theoretical research this is of limited importance provided the categorisation reveals interesting 'kinds' in nature
- However, if we wish to explore the existence of unconscious knowledge in given contexts then bias poses an issue

Potential Bias in Confidence Reports



Potential Bias in Confidence Reports



APPEAL TO SELF INTEREST

Gambling Paradigms and Incentives

- Pit self interest against any tendency to under report confidence
- Make clear the experimenters desire for the participant to be as accurate as possible

GAMBLING PARADIGMS - THEIR DRAWBACKS

Post Decision Wagering (Persuad et al., 2007)

- You bet high or low on each decision
- The relation between wager and accuracy is taken indicate metacognition

GAMBLING PARADIGMS - THEIR DRAWBACKS

Post Decision Wagering (Persuad et al., 2007)

- You bet high or low on each decision
- The relation between wager and accuracy is taken indicate metacognition
- *Is subject to risk aversion, r = .52 (Dienes & Seth, 2010)*

GAMBLING PARADIGMS - THEIR DRAWBACKS

Post Decision Wagering (Persuad et al., 2007)

- You bet high or low on each decision
- The relation between wager and accuracy is taken indicate metacognition
- *Is subject to risk aversion, r = .52 (Dienes & Seth, 2010)*

No loss Gambling (Dienes & Seth, 2010)

- You either bet on your decision or on a coin toss
- Willingness to bet on the coin toss over your decision is taken to indicate the absence of conscious knowledge
- Not subject to risk aversion

GAMBLING PARADIGMS - THEIR DRAWBACKS

Post Decision Wagering (Persuad et al., 2007)

- You bet high or low on each decision
- The relation between wager and accuracy is taken indicate metacognition
- *Is subject to risk aversion, r = .52 (Dienes & Seth, 2010)*

No loss Gambling (Dienes & Seth, 2010)

- You either bet on your decision or on a coin toss
- Willingness to bet on the coin toss over your decision is taken to indicate the absence of conscious knowledge
- Not subject to risk aversion
- *No genuine incentive for accurate reporting of guesses*

GAMBLING PARADIGMS - THEIR DRAWBACKS

Post Decision Wagering (Persuad et al., 2007)

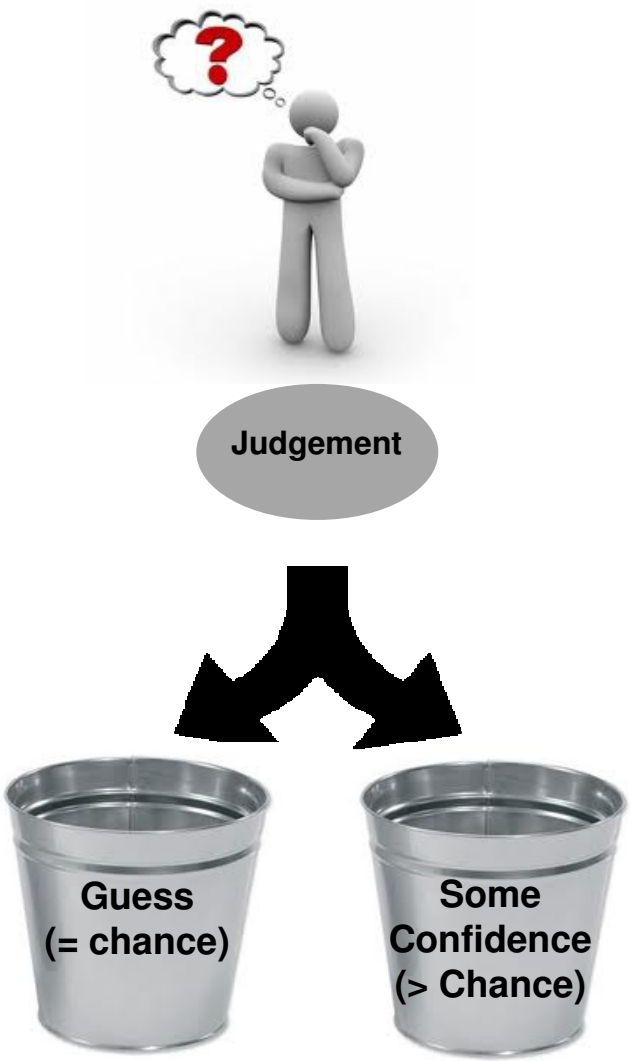
- You bet high or low on each decision
- The relation between wager and accuracy is taken indicate metacognition
- *Is subject to risk aversion, r = .52 (Dienes & Seth, 2010)*

No loss Gambling (Dienes & Seth, 2010)

- You either bet on your decision or on a coin toss
- Willingness to bet on the coin toss over your decision is taken to indicate the absence of conscious knowledge
- Not subject to risk aversion
- *No genuine incentive for accurate reporting of guesses*

We want a method that provides a genuine incentive and is not subject to risk aversion

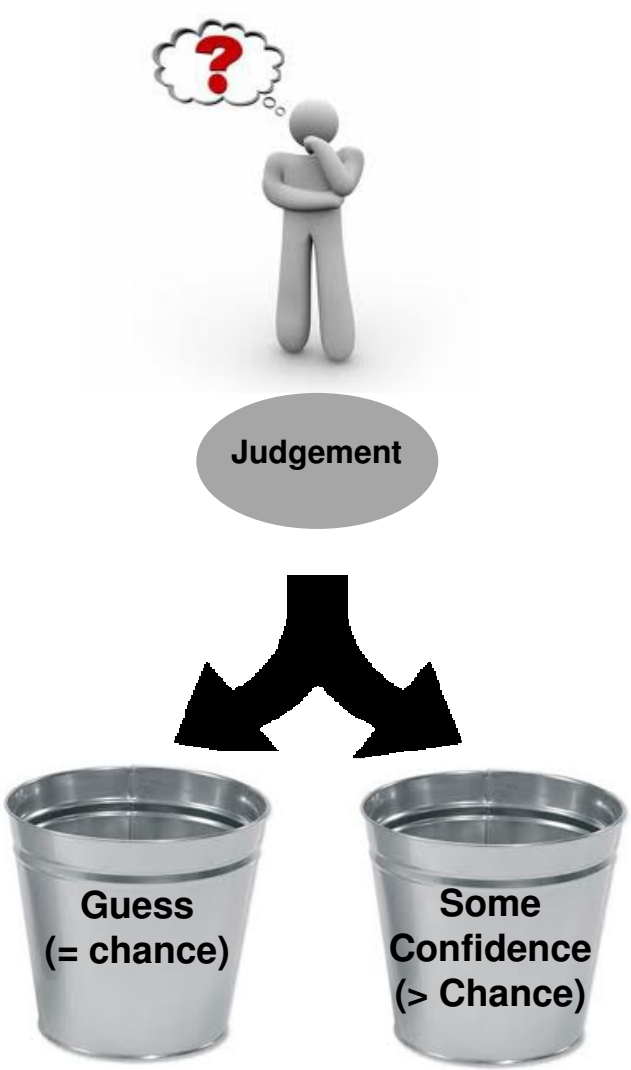
SYMMETRICAL CONFIDENCE INCENTIVES (SCI)



Method

- You start with a maximum payout
- You allocate each judgment to either:
  - ***Guess*** (at chance)
  - ***Some Confidence*** (above chance)
- A fixed amount is deducted for every wrongly allocated answer in **<u>either</u>** category (a symmetrical incentive)

SYMMETRICAL CONFIDENCE INCENTIVES (SCI)



Method

- You start with a maximum payout
- You allocate each judgment to either:
  - ***Guess*** (at chance)
  - ***Some Confidence*** (above chance)
- A fixed amount is deducted for every wrongly allocated answer in **<u>either</u>** category (a symmetrical incentive)

Characteristics

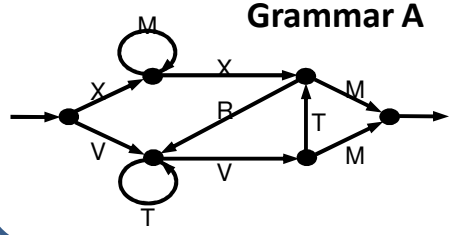
- Potential loss can't be varied or avoided so it is not subject to risk aversion
- Genuine incentive for accurate reports
- Maximum return achieved by reporting confidence as accurately as you can

ASSESSING SYMMETRICAL CONFIDENCE INCENTIVES

120 participants complete a standard artificial grammar learning task

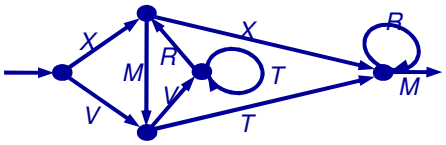
Trained on either

Grammar A



XMMXM
VTTVTM
VVTRTVM
…

Grammar B



XMTRM
VVRMTM
VTRRRRM
…

Tested on both

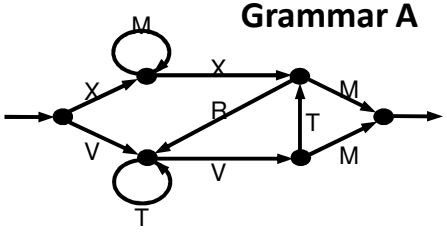
VTVTM
VTRRM
XXRVM
XXRRM
…

Assessing Symmetrical Confidence Incentives

120 participants complete a standard artificial grammar learning task

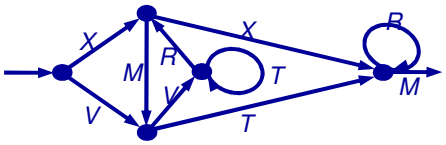
Trained on either

Grammar A



XMMXM
VTTVTM
VVTRTVM
…

Grammar B



XMTRM
VVRMTM
VTRRRRM
…

Tested on both

VTVTM
VTRRM
XXRVM
XXRRM
…

Judgements

Grammaticality judgment

Does the string obey the rules of the grammar?

Yes vs. No

Binary Confidence judgment

Do you have any confidence in your judgement?

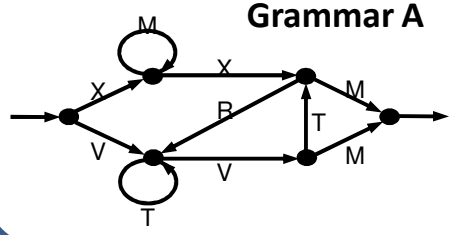
Yes vs. No Standard Confidence Reports	OR	Yes vs. No Symmetrical Confidence Incentives

Assessing Symmetrical Confidence Incentives

120 participants complete a standard artificial grammar learning task

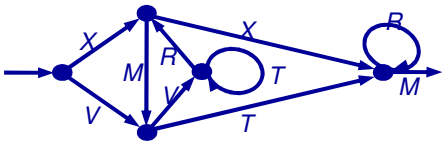
Trained on either

Grammar A



XMMXM
VTTVTM
VVTRTVM
…

Grammar B



XMTRM
VVRMTM
VTRRRRM
…

Tested on both

VTVTM
VTRRM
XXRVM
XXRRM
…

Judgements

Grammaticality judgment

Does the string obey the rules of the grammar?

Yes vs. No

Binary Confidence judgment

Do you have any confidence in your judgement?

Yes vs. No Standard Confidence Reports

OR

Yes vs. No Symmetrical Confidence Incentives

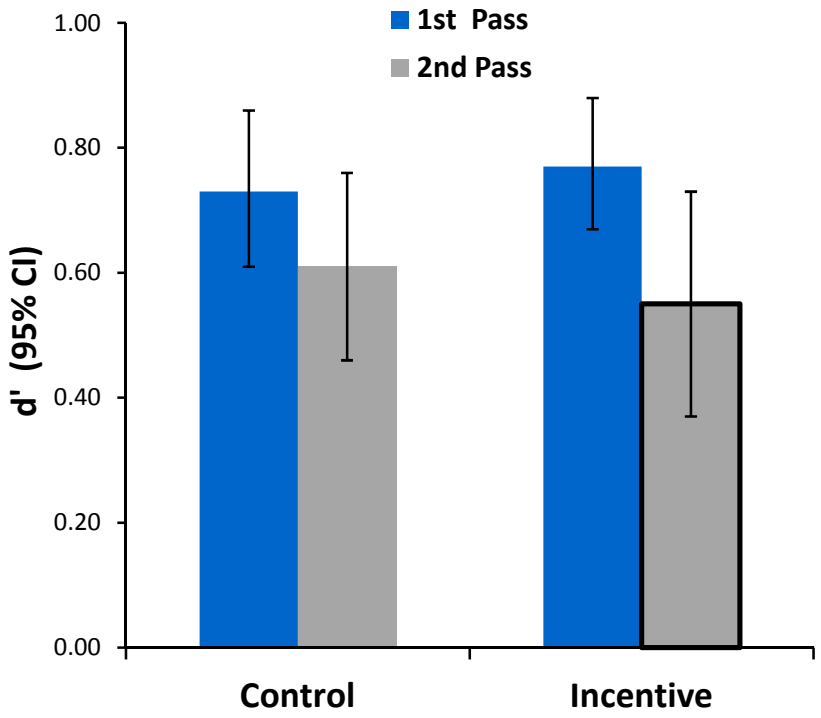
Design

Test Phase

	Pass 1	Pass 2
Control Condition	Standard Confidence Reports	Standard Confidence Reports
Incentive Condition	Standard Confidence Reports	***Symmetrical Confidence Incentives***

RESULTS

Classification performance

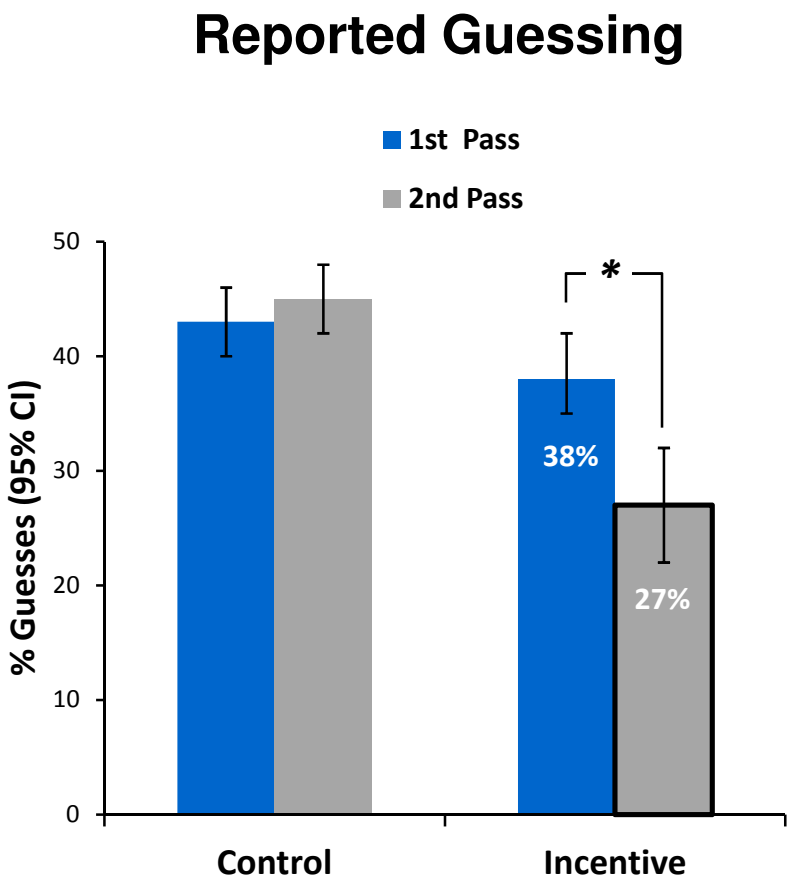


Classification performance was unaffected by the incentives.

Main effect of pass only
($F = 22.96$, $p < .001$)

Performance deteriorates over time irrespective of incentives.

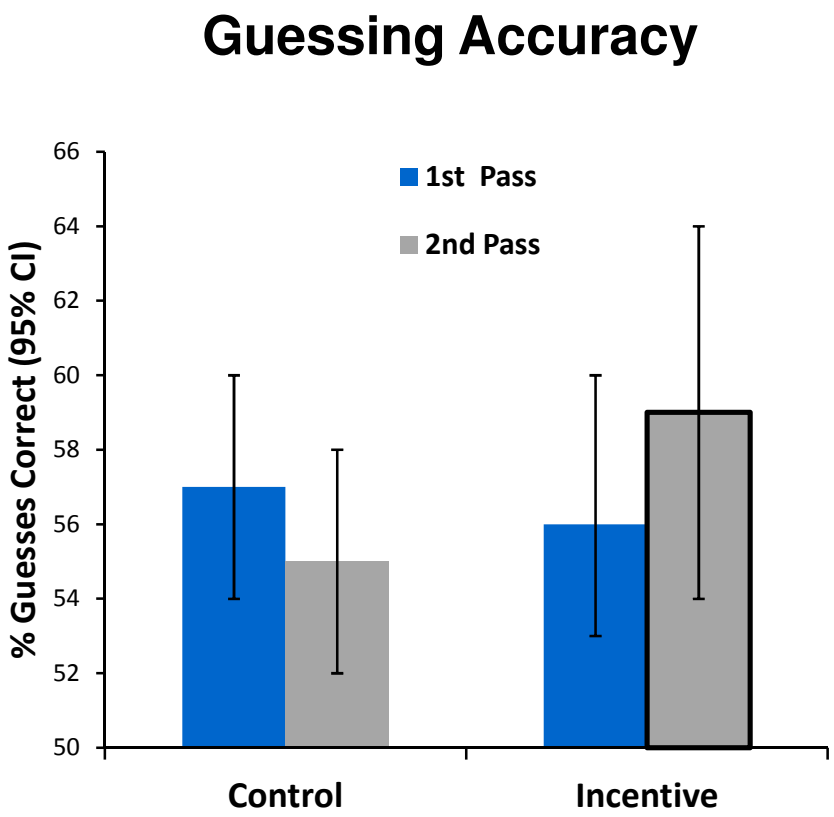
RESULTS



Proportion of guesses was significantly reduced by incentives.

Pass x Condition interaction
($F = 15.65$, $p < .001$)

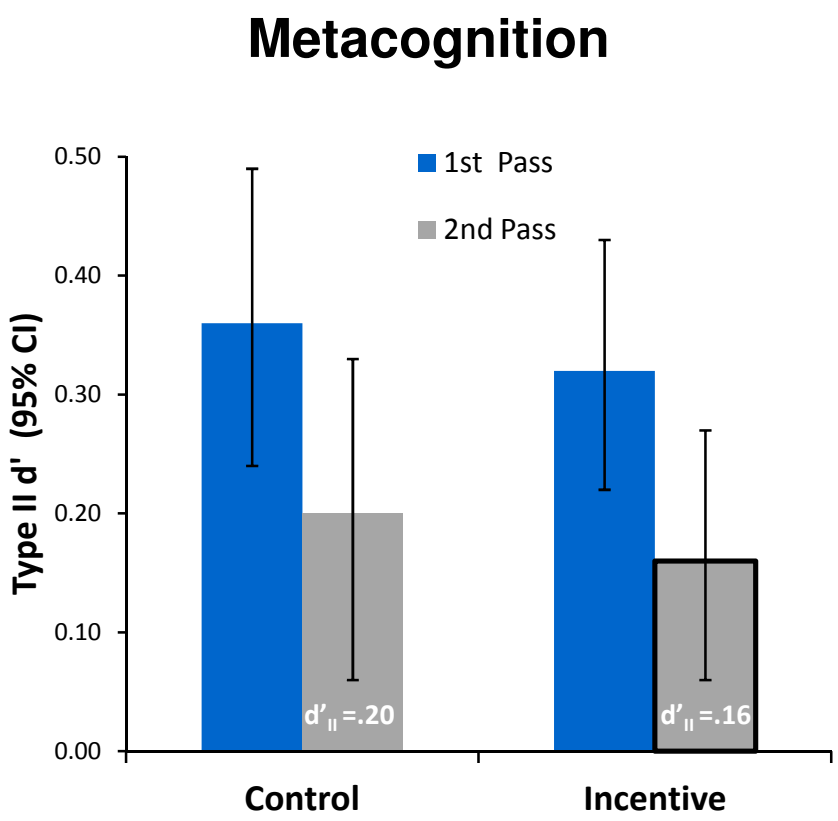
RESULTS



Accuracy of guesses was not significantly reduced by incentives.

95% CI excludes a reduction greater than 2%, which would still leave accuracy significantly above chance.

RESULTS

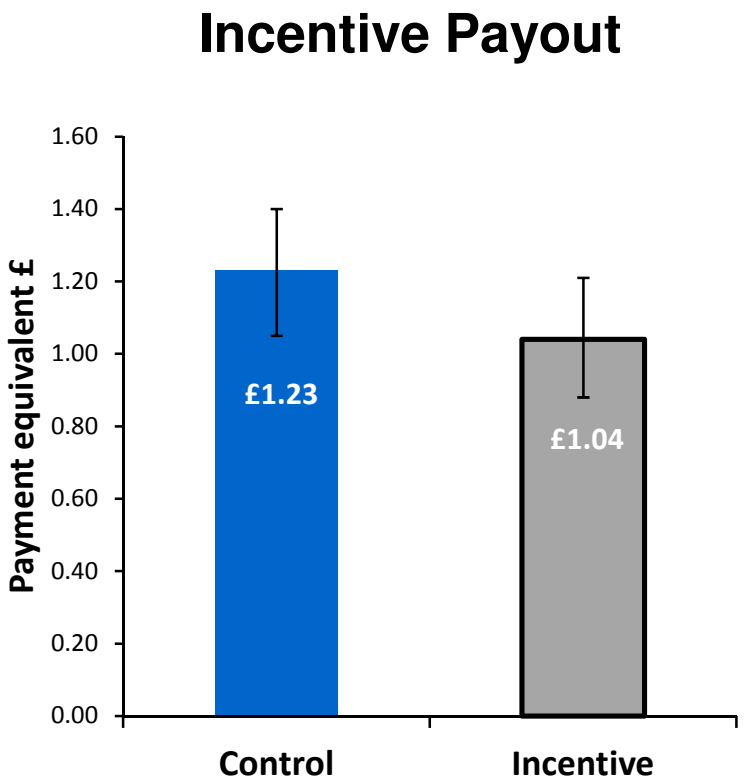


Metacognitive accuracy was unaffected by incentives.

Main effect of pass only ($F = 11.85$, $p = .001$)

Performance deteriorates over time irrespective of incentives.

RESULTS



Self interest was unable to increase the accuracy of confidence reports in AGL

*Participants would have earned 18% **more** without knowing about the incentives!*

CONCLUSIONS

- **Symmetrical Confidence Incentives did motivate people to try to report confidence more accurately – as evidenced by reduced guessing**

CONCLUSIONS

- Symmetrical Confidence Incentives did motivate people to try to report confidence more accurately – as evidenced by reduced guessing
- **In AGL that motivation did not reveal greater metacognitive insight than apparent using standard confidence reports**

CONCLUSIONS

- Symmetrical Confidence Incentives did motivate people to try to report confidence more accurately – as evidenced by reduced guessing
- In AGL that motivation did not reveal greater metacognitive insight than apparent in standard confidence reports
- **At least in the AGL paradigm standard confidence reports do not appear to underestimate metacognition**

CONCLUSIONS

- Symmetrical Confidence Incentives did motivate people to try to report confidence more accurately – as evidenced by reduced guessing
- In AGL that motivation did not reveal greater metacognitive insight than apparent in standard confidence reports
- At least in the AGL paradigm standard confidence reports do not appear to underestimate metacognition
- **We cannot assume this to be true of other paradigms but can employ Symmetrical Confidence Incentives to test them**

CONCLUSIONS

- Symmetrical Confidence Incentives did motivate people to try to report confidence more accurately – as evidenced by reduced guessing
- In AGL that motivation did not reveal greater metacognitive insight than apparent in standard confidence reports
- At least in the AGL paradigm standard confidence reports do not appear to underestimate metacognition
- We cannot assume this to be true of other paradigms but can employ Symmetrical Confidence Incentives to test them
- **More generally we can apply Symmetrical Confidence Incentives wherever we want a more defensible measure of metacognition.**

THANK YOU

CONFIDENCE

Not just a measure of metacognition…
The key to great acts of stupidity!