computational mechanisms underlying moral choices’ adjustment to descriptive norms

Philip Pärnamets*, Björn Lindström*, Simon Jangard & Andreas Olsson
philip.parnamets@ki.se
http://emotionlab.se

background & aims
• human moral behaviour follows not only moral preferences and relevant prescriptive norms but are also determined by social context and the prevalent descriptive norms [1]
• recent findings have shown that participants moral ratings of selfish/pro-social behaviours in a public goods game are affected by the frequency of that behavior in the player group [2]
• using computational modelling it is possible to decompose task performance onto latent variables mapping to psychological constructs
• the drift-diffusion model is a widely applied framework for understanding decisions, which has recently been applied to moral choices [3]

1) can previous findings on ratings be extended to binary moral judgments?
2) how does the frequency effect translate into computations in a DDM framework?

results - behaviour
• sensitivity to frequency of prosocial behaviour
• framing effect for what behaviour is being asked about: choices when asked about selfish behaviours being more responsive to descriptive norms

results - modelling
• model captures choice response time patterns well
• asking about selfish behaviours lowers decision boundaries (a*ASK)
• starting point bias (z) toward prosocial choice
• drift rate (d) sensitive to both amount of displayed behaviours and behaviour asked

method
98 participants recruited through Mturk
procedure
• participants are thoroughly acquainted with the procedure of public goods games (PGG).
• on each trial shown eight agents, identified with non-recurring letter combinations, who are said to be participants from earlier one-shot PGG experiments.
• under each participant their choice in a PGG is displayed: either keep or invest (selfish, prosocial respectively).
• one agent is selected and the participant is asked if she considers the selected behaviour to be morally right or wrong.
• 60 trials per participant

conclusions
• moral judgments about pro/antisocial behaviour is contextually sensitive to descriptive norms
• this sensitivity is a complex response captured by drift diffusion model
• selfish approving choices more error prone due to lowered boundaries and changed evidence sensitivity

References: