The hypothesis: Animacy perception (AP) relies on multiple cues.

* By analogy: Distance perception doesn’t just rely on a single cue, it relies on multiple (e.g., retinal disparity, motion parallax, vergence, etc.) and weighs these up (perhaps in a Bayes optimal manner).
* Similar points apply to other aspects of vision (e.g., size constancy). As such, it would be remarkable if AP simply relied on one cue.

Some of the most compelling evidence for AP: Brian Scholl’s ‘Wolfpack’ Paradigm.

Several features of this paradigm are notable:

* It has provided evidence that agents can be identified simply on the basis of *how they move*, and irrespective of how they look.
* Such identification appears to be judgement independent (you can know and reflect on the fact that the ‘agents’ aren’t actually agents), irresistible (it persists even when you’d like it not to), it highlights a compelling phenomenological contrast (agents and non-agents seem to visually appear different), and it appears to be the result of visual brain areas (personally, I’m not that moved by this last point). Collectively, this suggests that the agency identification in question is genuinely perceptual.

But, of course, movement is just one cue to animacy. The other (obvious) cue would be how things look. (Scholl et al. specifically want to eliminate this in their experiments, so it’s not that they’ve overlooked this point – it’s rather that their interest is in establishing AP)

Other work has investigated influences of how things look on AP, but (I think) the question of how it relates to the above remains un(der)explored and unanswered.

* Tremoulet & Feldman (2000) manipulated the shape/orientation of an ‘agent’ in a Heider & Simmel type task. They found that if the shape did not remain oriented along the direction of motion, subjects would be significantly less likely to report seeing it as an agent.
  + Problems: they simply asked subjects to rate levels of perceived animacy. So, it’s not clear that they tapped genuine AP as opposed to post-perceptual judgement. And this is clearly a problem: findings that things which ‘look’ more like agents [in something like Martin’s sense] are reported to look more like agents is kind of unsurprising. The interesting question is whether this is genuinely perceptual, in something like the way the wolfpack seems to be.
* Hernik, Fearon & Csibra (2014) highlight loads of reasons to think that anteroposterior body structure and symmetricality would be used to constrain both who is perceived as an agent, and what agents are going to do.
  + Their study probed the question of whether agency detection could be used to infer features of anteroposterior body structure (e.g., front from back) and rely on it when anticipating behaviour (it subsequently took longer for infants to stop looking at stimuli when they were ‘facing’ a target they had previously approached, than when they were ‘facing’ away from it).
  + But note: again, it’s not clear that this really reflects AP as opposed to infants’ judgements.
  + Furthermore, while the intro emphasises the importance of symetricality, this isn’t actually manipulated.
  + Finally, this doesn’t bear directly on the aforementioned hypothesis. In the case of depth perception, it’s not (just?) that vergence might enable some sub-module to better assess levels of motion parallax, or something like that. It’s rather that the visual system seems to weigh up information from these myriad sources (myriad depth cues) and uses these to (flexibly) infer the most likely depth hypothesis (perhaps in a Bayes optimal way – Ernst & Banks, 2002; Rescorla 2015; Trommershäuser et al. 2011; cf. Orlandi 2014). So, while suggestive, the key prediction of my hypothesis remains untested. It’s not enough to show that we/infants can identify agents, and then on the basis of that identify other features of agents.

So, here’s my suggestion:

Experiment 1, replicates a prominent Wolfpack study (e.g., Gao et al., 2009).

* Observers watch simple shapes move around a computer display, with one shape (the wolf) pursuing another (the sheep).
* In the *Don’t-Get-Caught* task, subjects control the sheep’s movements by moving a computer mouse, attempting to avoid the wolf. But, because there are so many objects in the array, they can’t avoid them all. And while the sheep is always highlighted, the wolf is not (it looks identical to/indistinguishable from distractors). As such, subjects have to detect chasing *entirely on the basis of it’s spatiotemporal movements*.
* Following Gao et al., we vary chasing subtlety (the wolf’s angular deviation from perfect heat seeking). 0degrees and it heads straight for the sheep, 90degrees and it heads completely orthogonal to the sheep. And when deviation is 30degrees, it heads towards the sheep within a 60degree arc of perfect heat seeking.
* The result (if it replicates) is that 0degrees is easy to detect (it’s very obvious that the wolf is the object going directly for you) and subjects readily avoid the wolf (I guess you need to ensure that the wolf doesn’t move too quickly). And if it’s 180degrees then it’s just incompetent, so easily avoided. But there’s a U-shaped drop in performance in between. So we replicate that and measure the U-shaped curve.

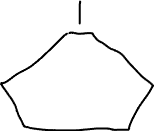
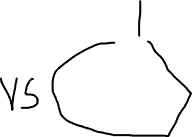
Experiment 2 is the same as the above, but we now vary symmetricality of the shapes. So, now, some shapes are symmetrical and some aren’t. Low-level confounds (like total area) are controlled for (we don’t want agents to be bigger on the screen, or otherwise more attention grabbing/consuming).

* If symmetricality is a cue used for animacy detection, it should be easier to detect a symmetrical wolf at the bottom of the aforementioned U-shape curve. So, if 60degree chasing subtlety is maximally dangerous in the first experiment, we would expect that <60degree chasing subtlety will be more dangerous when the wolf is asymmetric.
  + Notice that the assymetrical wolf condition is quite different from Tremoulet & Feldman, where (really) consistent orientation as opposed to assymetricality was manipulated:

Shape, circle

Description automatically generated with medium confidence

Manipulation in Tremoulet & Scholl.



Crap drawing of an asymmetrical vs. symmetrical wolf. (Maybe we’d need to/should keep other low-level features like no. of sides)

I think that the results would be interesting either way.

* If our prediction is borne out, then this indicates that AP (and not just post-perceptual judgement) is influenced by myriad cues and involves a weighting of these.
* If our prediction is not borne out, then it indicates that AP is (relatively) encapsulated from biologically salient information which others expect to influence AP proper.