Developmental Science 1:1 pp 71–78

REPORT

Computational evidence for the foundations of numerical competence

Tony J. Simon

Georgia Institute of Technology, USA

Abstract

INFANT is a computational model of Simon's (1997) 'non-numerical' account of the foundations of numerical competence. It operationalizes the memory, individuation, object permanence and spatiotemporal representation competencies of that account into a simulation of the Simon, Hespos & Rochat (1995) study. Results demonstrate that infants' responses can indeed be reproduced without numerical representations.

Wynn (1992) claimed that 'Humans innately possess the capacity to perform simple arithmetical calculations' and that '[they] possess true numerical concepts – they have access to the ordering of numerical relationships between small numbers' (p. 750). However, no information processing account of such ability has been presented. In contrast, Simon (1997) presented a detailed alternative positing that, instead of numerical representations and processes, a small set of largely domain-general infant competencies is sufficient to produce the observed behavior. That account forms the basis for the computational model presented here. The model faithfully reproduces the responses of babies in a published study (Simon, Hespos & Rochat, 1995). Furthermore it supports the claim that infants in that study responded to physical impossibility on the basis of spatiotemporal object representations and not on the basis of numerical knowledge.

The general task used in these studies can be described as follows (see Figure 1). Five-month-old infants see a puppet stage upon which sits a small doll. After an infant's attention has been drawn to the doll it is occluded by raising a screen. Another doll is shown beside the stage and then placed behind the screen. Next, the infant sees the experimenter's empty hand leave the stage area. The screen is then lowered, revealing some or all of the dolls. The primary dependent variable is the infant's looking time upon lowering the screen. Longer looking times are interpreted as a violation of the infant's expected state of the world and are assumed to indicate some kind of examination in order to comprehend the outcome. Here, the unexpected outcome is when only a single doll remains on the stage. (Studies typically include the inverse task where both dolls are initially present, one is removed and either a single doll or both dolls are revealed).

Several studies have shown that infants look longer at the unexpected than the expected outcomes. Counter to Wynn's (1992) claim that this response is based on numerical knowledge, Simon's (1997) position is that the basis is an appreciation that the outcomes violate object permanence laws. However, Simon *et al.* (1995) were unable to induce expectation-violations in 'identity' conditions where the correct number of objects was revealed but where the physically-impossible transformation of one object's identity into another had occurred. Wynn (1995) interpreted that result as support for her numerical interpretation. Simon claims the correct explanation is that infants did not use the identity of the objects involved when evaluating the outcome. This view is based on Xu and Carey's (1996) finding that, in similar tasks, infants prefer a spatiotemporal encoding of objects to one based on identity. The model presented here provides evidence that non-numerical spatiotemporal representations are a sufficient basis for infants' responses.

Simon (1997) specified four areas of competence sufficient to generate the observed behaviors. One is the

Address for correspondence: Tony J. Simon, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170, USA
email: tony.simon@psych.gatech.edu

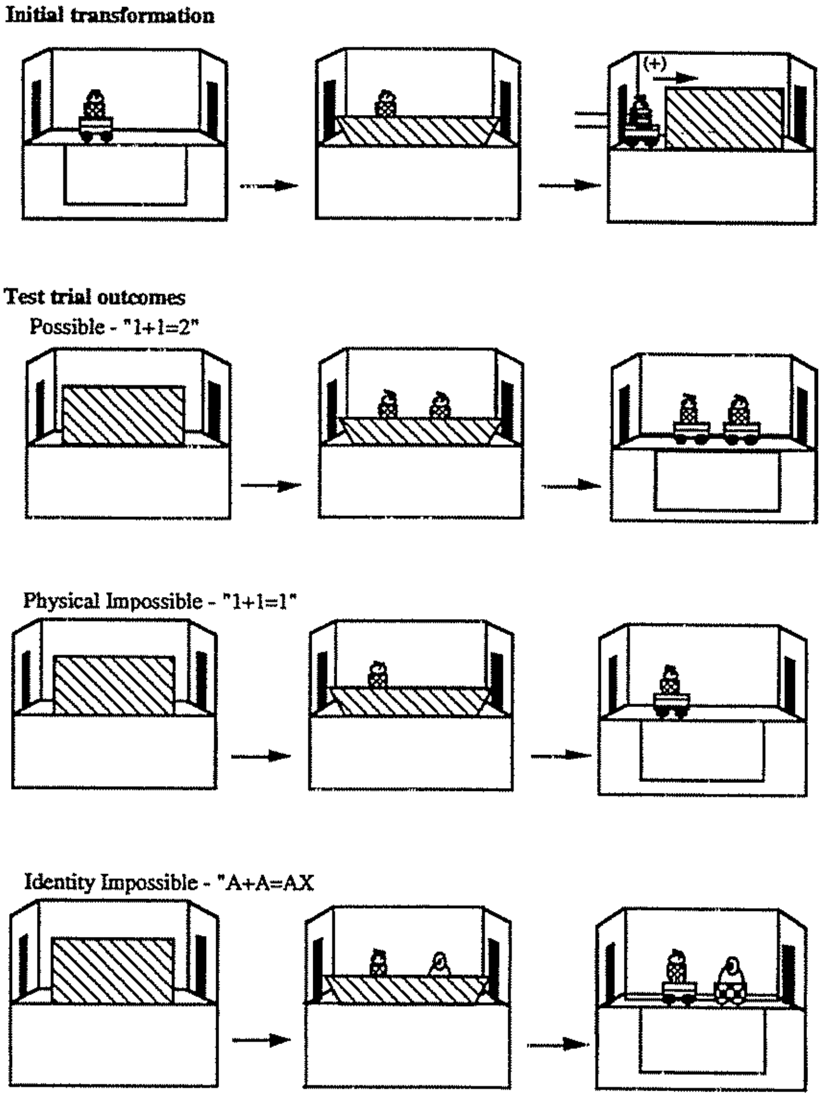

Figure 1 *Sequence of events for the '1 + 1' or 'addition' tasks.*

memory and discrimination abilities which enable infants to compare remembered and visibly presented entities. These are the minimal requirements for habituation, a primary infancy research tool. Individuation competence adds constraints by limiting, to around three, the set of items that can be represented as unique entities. Such competence is evident in infancy (e.g. Antell and Keating, 1983), and is likely to be explained in terms of early visual attention processes (Trick and Pylyshyn, 1994). To generate expectation-violations in these experiments, infants must represent the persistence of occluded objects. Object permanence, the third area of required competence, has been demonstrated in infants at 2 months of age and beyond (e.g. Baillargeon, 1993). Finally, infants must represent objects primarily in terms of spatiotemporal characteristics. Simon also claimed that outcome evaluation is based on a simple one to one matching algorithm.

General description of INFANT

Following the above requirements, I have constructed a computational model, called INFANT, which simulates an infant presented with the tasks described by Simon *et al.* (1995). Though the tasks are labeled with arithmetical terms (see Table 1) this does not indicate that I consider numerical competence to be required. Beyond the theoretical assumptions described in Simon (1997), INFANT inherits many key aspects of the theory of the

Table 1 *Label and description of all tasks presented to INFANT.*

Task	Initial state	Action	Result state
1 + 1 = 2	Elmo	Introduce Elmo	Elmo and Elmo
1 + 1 = 1	Elmo	Introduce Elmo	Elmo
2 − 1 = 1	Elmo and Elmo	Remove Elmo	Elmo
2 − 1 = 2	Elmo and Elmo	Remove Elmo	Elmo and Elmo
A + A = AX	Elmo	Introduce Elmo	Elmo and Ernie
AA − A = X	Elmo and Elmo	Remove Elmo	Ernie

ACT-R cognitive architecture (Anderson, 1993) in which it was constructed (see Simon and Halford, 1995 for a discussion of the theoretical role of architectures). Competence is encoded in ACT-R in the form of production or if-then rules. Each rule is evaluated by ACT-R as a candidate for firing, based on a computation of its expected gain (or progress towards the goal) minus its expected cost (of achieving the goal by that means). The activation level in memory of the elements tested by a rule affects the time that rule takes to match[1] (see Anderson, 1993, Chapter 3). This last feature allows dynamically-varying match times to be computed by ACT-R. These are added to the rule's execution time to produce latency measures for individual actions and the entire task. Since simulated chronometric values such as looking time, in the case of INFANT, are produced, ACT-R supports quantitative as well as qualitative simulations of mental activity.

Representational conventions

To capture the key characteristics of an experimental environment like Simon *et al.*'s, INFANT operates in a simplified world compartmentalized into three positions: Left, Center and Right.[2] Objects occupying the Left and Center positions are ones placed on the stage. These can be visible or hidden, depending on whether or not the screen is raised. Objects in the Right position are not on the stage and are in the process of being placed on, or removed from the stage, and thus are always visible (Figure 1 depicts these positions reversed). To afford a spatiotemporal representation of the task, it is broken into a sequence of distinct temporal frames. No absolute time is assigned to these. Instead the frame is described as 'advanced' when an experimenter-induced event, like placing an object on the stage, occurs. Once all experimenter actions for an event are completed, simulations of the baby's mental actions begin. Their completion signals the end of a temporal frame, which is reset to the value 'static'. Thus, in a typical temporal frame (e.g. Frame 1 in Figure 2) the baby is presented with a visible object on the stage and, in response, creates an encoding of it.

Physical objects, i.e. ones that exist in the real world, are coarsely described with 5 attributes: Location, Motion, Support, Status and Token. Location can take the values [Left, Center, Right], and Motion can take the values [Leftward, Rightward, None], depending on where the object is and its direction, if any. Support can take the values [Stage, Hand], determined by whether or not the object is being held. Status can take the values [Visible, Hidden, Removed]. Status [Removed] denotes that an object is remembered as having been recently present. Finally, Token can take the values [Elmo, Ernie], depending on which doll is being referred to. Notice that Token values contain no descriptions of the dolls and could easily be replaced by arbitrary values like 'foo' and 'bar'. Their only purpose is to identify the actual token involved for tasks such as A + A = AX. Thus, an Elmo doll in the process of being placed on the stage, would be described as: Location [Right], Motion [Leftward], Support [Hand], Status [Visible], Token [Elmo].

Objects maintained in the infant's memory are encoded even more coarsely, with just abstract existential descriptors. Each object is represented by an 'Object File' index (e.g. Gordon and Irwin, 1996) which can have up to 3 attributes. Type and Object are obligatory, with Prediction being the optional attribute. Type takes the values [Phys-obj, Mem-obj]. Phys-obj refers to a physical object that can be seen, i.e. Status [Visible], while Mem-obj is one that is retained in memory, i.e. Status [Hidden or Removed]. Object is a link to a physical object. This is required for 2 reasons: To ensure each physical object has a unique index, and to enable predictions to access the spatiotemporal description of the object via this link. Thus, to directly implement Simon's (1997) theory, Predictions inherit the existing values of a physical object only for the attributes Location, Motion and Support. For the tasks being modeled here, predictions only arise for objects that are static. However, predictions of the future locations of moving objects could be included.

---

[1] A number of parameters in ACT-R can affect the dynamically-varying activation levels. In INFANT the only non-default parameter involved is Base Level Learning (bll), and this is set to 0.25 for all tasks.

[2] This currently has effect of implying that infants know precisely where an object is behind the screen at any moment, which conflicts with some recent evidence (e.g. Koechlin, Dehaene and Mehler, in press). This implication could easily be removed by renaming the locations 'onstage1', 'onstage2' and 'offstage' without affecting the behavior of the model in any way.

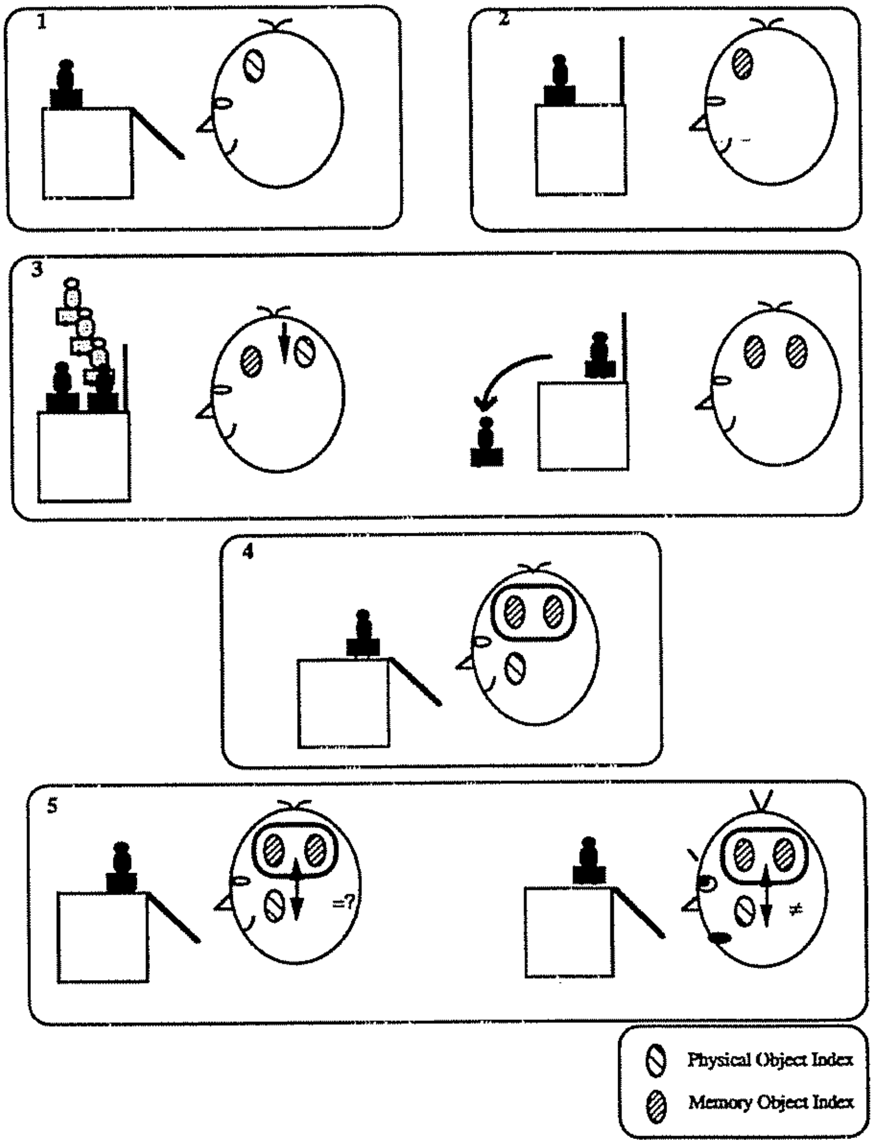

Figure 2 *Processing diagram depicting representations by temporal frames.*

INFANT's processing in the '1 + 1 = 1' task

Here I explain how INFANT's competence generates behavior in response to the task depicted in the temporal frames of Figure 2. Processing variations caused by other tasks will be discussed in context. Whenever the screen is lowered and INFANT has no Phys-obj indexes, it creates an index for each visible object. That is what occurs in Frame 1. Frame 2 shows the same scene but with the screen now raised. Whenever INFANT has a Phys-obj index for an object whose status is changed from Visible to Hidden, as is the case now, it changes the index to Type [Mem-obj]. Since all Mem-obj indexes carry predictions, the physical object's spatiotemporal characteristics are copied to the index. INFANT now has an expectation for an object with Location[Left], Motion [None], Support [Stage]. In Frame 3 a new object is presented visibly, and then hidden behind the screen. INFANT creates a new Phys-obj index for that object, as above, and then changes it to Type [Mem-obj] once it is hidden, and adds the prediction. It now has a second expectation; for an object with Location [Center], Motion [None], Support [Stage]. In the second part of Frame 3 the experimenter removes one of the objects. Since INFANT does not witness this change no change is made to the indexes and their predictions.

In Frame 4, the screen is lowered revealing (in this case) one physical object. Since INFANT has no indexes of Type [Phys-obj] it again creates them, resulting in a single index whose spatiotemporal characteristics are Location [Center], Motion [None], Support [Stage]. Now INFANT is in a state where it has newly-created physical object indexes and pre-existing memory objects which carry predictions for an expected state of the world. In such a state, INFANT creates a subgoal, called Verify-Prediction, whose purpose is to compare the predicted to the actual state of the world. The selection of that subgoal signals the start of the looking time measure, and the termination of the subgoal with Value [Match/Mismatch] signals the end. Within this subgoal, INFANT carries out a simple one to one match process as often as is necessary, proceeding as follows. First INFANT selects one of its expected objects (i.e. Type [Mem-obj]) and attempts to match its prediction to one of the physical objects on the stage (Frame 5). Due to the ease of matching correct predictions INFANT verifies these first. Since this match can be carried out by a single rule, it takes very little processing time to execute. The rule simply matches the Location, Support and Motion values of a predicted object with those of a physical object with the same values. When a match occurs the verification process is re-initiated with any other predictions that exist. If no matching object can be found, then one of two search procedures is triggered.

Frame 5 indicates a match between the expected and physical objects with Location [Center], Motion [None], Support [Stage] values. However, INFANT still has another expectation (specifically, Location [Left], Support [Stage], Motion [None]), but no physical object matches that description. Having found no match at Location [Left], INFANT searches in the two other locations to see if it can find a match for the predicted object. Since this will never be possible, it will terminate the subgoal with Value [Mismatch]. The extra actions required to check the locations against the prediction are what cause the prolonged processing time (taken to indicate looking time) in the subgoal.

INFANT's processing in other tasks

In a $2 - 1 = 2$ task, INFANT would begin by creating 2 indexes, change one to Status [Removed] for the removed object, and have just a single expected object in its memory. Now, one expectation will match a visible object but INFANT's memory will include an expectation for an object with Location [Right], Support [Hand], Motion [Rightward], Status [Removed], representing the details of the object that was removed. The index INFANT creates for the second physical object will not match that expectation's characteristics. INFANT determines that by checking the value of each attribute, one at a time, to see if there is a match between the removed and remaining physical object. Since these values will not match, the subgoal will again be terminated with Value[Mismatch].

Processing in the identity conditions (e.g. $A + A = AX$) proceeds exactly as above except that both expected objects will match visible objects. This is because the Token attribute is not tested in matching the expected and visible items. INFANT cares only that 'an object' is expected to exist, and not which object it is. However, INFANT does suggest that some notice may be taken of the change. Aggregate looking times in Simon *et al.* (1995) and those of INFANT both increase just slightly under these conditions. The ACT-R architecture has a clear theoretical reason for this longer latency. The new 'Ernie' token in $A + A = AX$, having had no processing history in the task, has a much lower activation level than the 'Elmo' that was present from the start. So, although its characteristics pass the test used by the matching rule, its lower activation creates a slightly longer latency. It is as if INFANT notices something has changed but does not consider it a big enough difference to invalidate its expectation. Perhaps babies too consider the identity change to be unimportant since their decision is based on spatiotemporal representations.

The veracity of INFANT's prediction-verification processes are unclear because it is not currently known what mental activities result from expectation-violations. However, INFANT's actions are similar to the perceptual analysis process proposed by Mandler (1988) as the basis for representation change in infancy. In order to balance INFANT's 'looking time' for different outcomes, the two search procedures each check 3 values for a match. This also seems suitable since 3 is the empirically-derived representational limit for sequential actions or events in infancy (e.g. Uller, Carey, Huntley-Fenner, and Klatt; 1996; Wynn, 1996). So while INFANT's expectation-violation processing may be qualitatively unlike that of babies, quantitatively speaking, INFANT's actions stand as a clear operationalization of the expectation-violation that is so central to infant research at present.

Results

INFANT's behavior on two 'addition' ($1 + 1 = 2$ and $1 + 1 = 1$) and 'subtraction' tasks ($2 - 1 = 1$ and $2\text{-}1 = 2$) was compared to the aggregate looking times of the infants in the Simon *et al.* (1995) study. The results are presented in Figure 3.

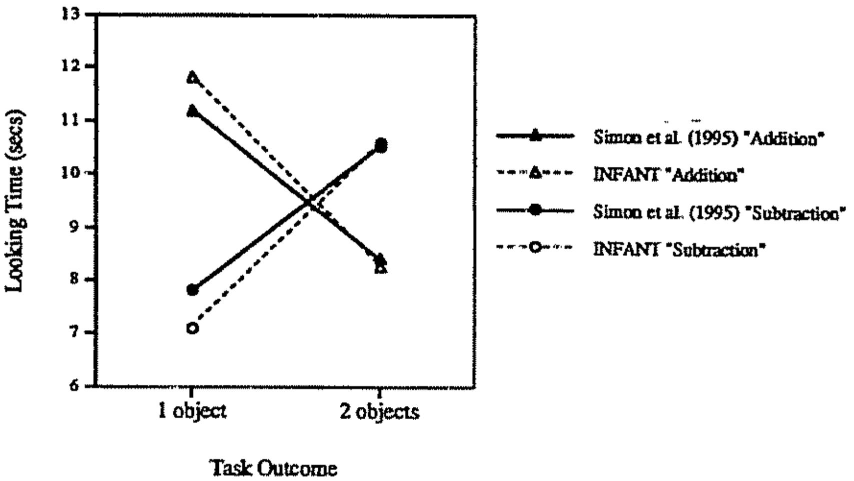

Figure 3 *Comparison of looking times for Simon* et al. *(1995). data and INFANT on standard 'addition' and 'subtraction' tasks.*

INFANT's behavior was qualitatively and quantitatively similar to babies in the Simon *et al.* (1995) study for these outcomes. Simulated looking time was longer for 'impossible' than for 'possible' outcomes because the latter could be verified by simple matches, while the former required deliberate search processing. This result is supporting evidence for INFANT; a detailed explanation of the competence that 5-month-olds employ in so-called infant number tasks. It further supports Simon's (1997) assertion that a system of cardinal and ordinal concepts and calculational ability is not required, because there exists no such competence in INFANT. It demonstrates that a set of domain-general abilities involving simple individuation, memory, spatiotemporal representation, and object permanence, is *sufficient* to generate the observed behavior.

However, the model's main role concerns the nature of representations underlying infants' responses in Simon et al.'s (1995) 'identity' conditions. Since INFANT contains no numerical representations it cannot evaluate outcomes like 'Elmo + Elmo = Elmo and Ernie' in numerical terms as Wynn claimed. It can only respond on the basis of physical possibility, or object permanence. Further, if INFANT processed identity and not spatiotemporal objects features its behavior would be inconsistent with the data. This is because a rule recognizing Elmo features would not fire when Elmo was replaced by an Ernie. A mismatch, rather than the observed match would result. Thus, to simulate these human data INFANT must represent objects so as to expect 'that "a thing and a thing behind the screen should lead to a thing and a thing behind the screen", which it did!' Simon (1997, p. 368). Not only did INFANT closely replicate the standard 'possible' and 'impossible' outcomes, it is evident from Figure 4 that INFANT's simulated looking times in the 'A + A = AX' and 'AA − A = X' conditions are almost exactly the same as the human infants. Thus, INFANT strongly supports the claim that an abstract representation of objects, and not number knowledge, is crucial to the human competence account.

INFANT also offers a novel explanation for why infants typically show longer looking times for 'possible' displays containing 2 objects than for single object displays (e.g. Simon *et al.*, 1995; Wynn, 1992; Xu and Carey, 1996). Rather than having a preference for 2 objects, INFANT simply takes longer to encode 2 objects than 1. This is because each visible object requires an individuation rule to fire and place an index in memory. Figure 5 (and Figure 3) shows that, in the absence of mismatched predictions, INFANT takes

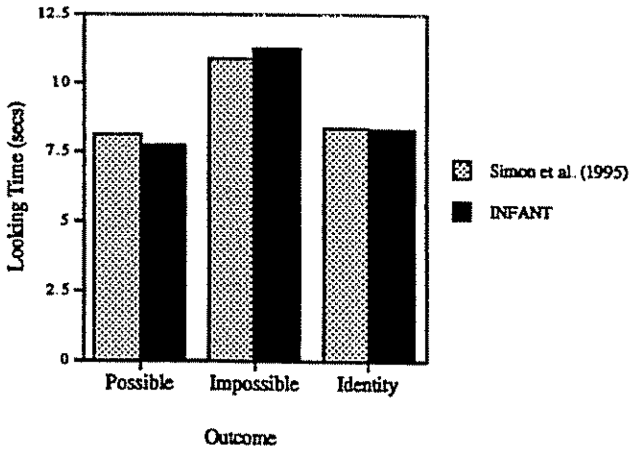

Figure 4 *Comparison of looking times for Simon* et al. *(1995) data and INFANT on possible, impossible and identity tasks.*

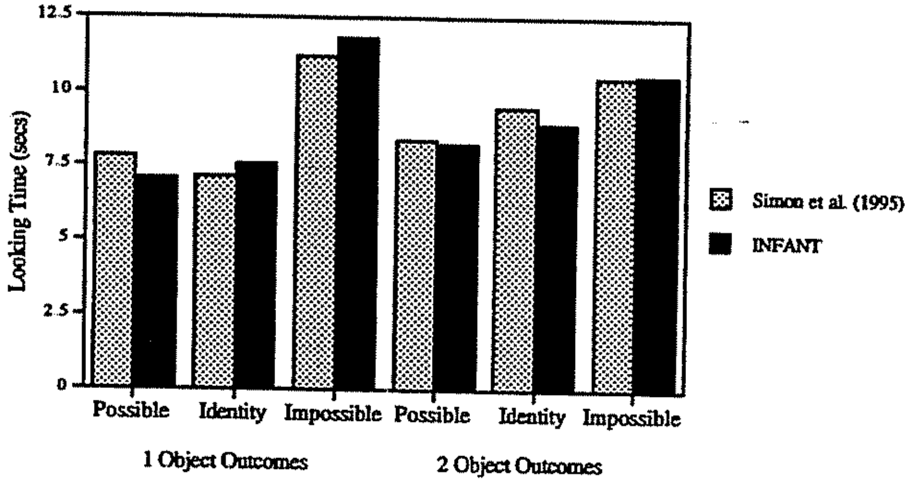

Figure 5 *Comparison of looking times for Simon* et al. *(1995). data and INFANT on all 1-object versus 2-object outcomes.*

longer to process 2-object displays than 1-object displays.

Conclusions

INFANT closely replicated data from Simon et al.'s (1995) study. Because it operationalizes Simon's (1997) competence specification it provides evidence that those abilities are sufficient to reproduce the observed responses. By reproducing human behavior without any recognizable numerical competence INFANT shows that there is no need for the numerical competence Wynn has claimed to be an innate endowment of the human infant. It is important to stress that, because INFANT *can* reproduce the responses of human infants, it does not follow that infants employ the exact representational and processing constructs specified by the model. INFANT merely demonstrates the sufficiency of those competencies. That means, however, that innate numerical competence is not necessary to explain the published results concerning numerically-relevant behavior in human infants.

Not only is innate numerical competence unnecessary, there is a growing view that it is not even possible. Recent neurobiological approaches to development have argued against the idea that an organ as complex and flexible as the human brain could have innately-specified representations (Elman *et al.*, 1996). By implication, the only sound scientific course for investigating infant abilities is to propose a set of foundational competencies and their computational bases. These must be accompanied by a transformational account of the construction of more mature conceptual states; such as the system of numerical competence found in school-aged children. By presenting INFANT, I have made a start in that process. It is by testing the hypotheses generated by the model that the understanding of the foundations and development of human numerical competence will be enhanced.

Acknowledgement

I would like to thank John Anderson for comments on an earlier draft of this paper.

References

Anderson, J.R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Antell, S.E., & Keating, D.P. (1983). Perception of numerical invariance in neonates. *Child Development*, 54, 695–701.

Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C.E. Granrud (Ed.) *Visual perception and cognition in infancy* (pp. 265–315). Hillsdale, NJ: Lawrence Erlbaum Associates.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Koechlin, E., Deheane, S., & Mehler, J. (in press) Numerical transformations in five-month-old human infants. *Mathematical Cognition.*

Gordon, R.D., & Irwin, D.E. (1996). What's in an object file? Evidence from priming studies. *Perception and Psychophysics*, 58, 1260–1277.

Mandler J. M. (1988). How to build a baby: On the development of an accessible representational system. *Cognitive Development*, 3, 113–136.

Simon, T.J. (1997). Reconceptualizing the origins of number knowledge: A 'non-numerical' account. *Cognitive Development*, **12**, 349–372.

Simon, T.J., & Halford, G.S. (1995). Computational models and cognitive change. In T.J. Simon and G.S. Halford (Eds.) *Developing cognitive competence: New approaches to process modeling* (pp. 1–30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, T.J., Hespos, S.J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, **10**, 253–269.

Trick, L.M., & Pylyshyn, Z.W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, **101**, 80–102.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1996). What representations might underlie infant numerical knowledge. Manuscript under review.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749–750.

Wynn, K. (1995). Origins of numerical knowledge. *Mathematical Cognition*, **1**, 35–60.

Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science*, **7**, 164–169.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, **30**, 111–153.

Received 8 September 1997
Accepted 8 October 1997