Cognition 166 (2017) 407–417

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Desirable difficulties during the development of active inquiry skills

George Kachergis [a,*], Marjorie Rhodes [b], Todd Gureckis [b]

[a] Artificial Intelligence Dept./Donders Institute for Brain, Cognition, and Behavior, Radboud University, Nijmegen, Netherlands
[b] Department of Psychology, New York University, New York, NY, United States

ARTICLE INFO

Article history:
Received 18 July 2016
Revised 12 May 2017
Accepted 14 May 2017
Available online 13 June 2017

Keywords:
Question asking
Information search
Active inquiry
Hypothesis testing
Scientific reasoning

ABSTRACT

This study explores developmental changes in the ability to ask informative questions, hypothesizing a link between the ability to update beliefs in light of evidence and the ability to ask informative questions. Five- to ten-year-old children played an iPad game asking them to identify a hidden insect. Learners could either ask about individual insects, or make a series of feature queries (e.g., "Does the hidden insect have antenna?") that could more efficiently narrow the hypothesis space. Critically, the task display either helped children integrate evidence with the hypothesis space or required them to perform this operation themselves. Our prediction was that assisting children with belief updating would help them formulate more informative queries. This assistance improved some aspects of children's active inquiry behavior; however, despite making some updating mistakes, children required to update their own beliefs asked questions that were more context-sensitive and thus informative. The results show how making a task more difficult can improve some aspects of children's active inquiry skills, thus illustrating a type of "desirable difficulty" for reasoning.

1. Introduction

A skill of central importance during development is learning how to ask informative questions in order to make sense of the world. The roots of these abilities are observable even in the early preschool years. For example, in simple causal reasoning tasks, preschool-aged children can distinguish confounded from unconfounded evidence to draw causal inferences (Gopnik, Sobel, Schulz, & Glymour, 2001; Kushnir & Gopnik, 2005, 2007; Schulz & Gopnik, 2004). Preschool-aged children also selectively explore confounded evidence in their own exploratory play (Cook, Goodman, & Schulz, 2011; Gweon & Schulz, 2008; Schulz & Bonawitz, 2007). Despite these early emerging abilities, many of the cognitive skills required for self-guided, active inquiry seem to follow protracted developmental trajectories. For example, in tasks designed to assess scientific reasoning abilities, children in the older elementary school years (ages 8–10) often have difficulty adopting systematic strategies, such as testing the effects of one variable at a time or selecting interventions that will lead to determinate evidence (Chen & Klahr, 1999). Although children in the older elementary school years can be taught to engage in these strategies via direct instruction (Klahr & Nigam, 2004; Kuhn & Dean, 2005), it is notable how difficult it is for them to discover and implement them on their own.

One reason for the difficulties children exhibit in these types of inquiry tasks may be that active inquiry depends on the coordination of a variety of component cognitive processes (Bonawitz & Griffiths, 2010; Coenen & Gureckis, 2015). For example, according to one popular view (Klein, Moon, & Hoffman, 2006a, 2006b; Russell, Stefik, Pirolli, & Card, 1993), active inquiry unfolds as a sequence of mental steps (see Fig. 1). Learners must generate possible hypotheses to explain their environment. They then must engage in decision making to ask questions or gather additional information to decide which of these hypotheses is most likely. They then must understand the results of these inquiry behaviors and update their beliefs accordingly, and so on. The various stages of this loop closely mirror the process of scientific reasoning engaged by scientists (Klein et al., 2006a, Klein, Moon, & Hoffman, 2006b; Russell et al., 1993). Inefficiencies in any or all of these interrelated processes may serve as developmental limitations. For example, young learners may be able to search efficiently for information given a particular set of hypotheses but have trouble updating their beliefs correctly given new evidence. In this sense active inquiry behavior is like a bicycle: when all the

---

* Corresponding author.

E-mail addresses: g.kachergis@donders.ru.nl (G. Kachergis), marjorie.rhodes@nyu.edu (M. Rhodes), todd.gureckis@nyu.edu (T. Gureckis).

http://dx.doi.org/10.1016/j.cognition.2017.05.021

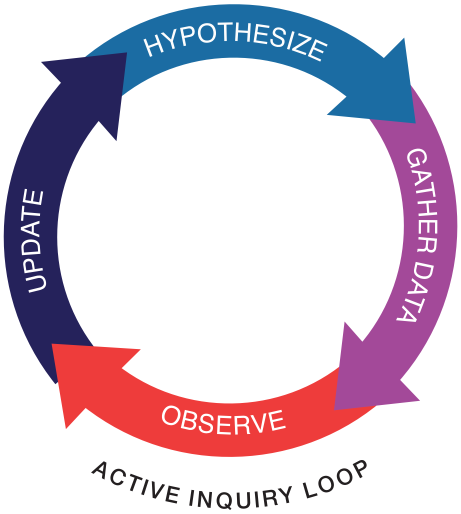

Fig. 1. The active sensemaking loop depicts the successive cognitive process that are engaged when attempting to derive a meaningful understanding of an initially ambiguous situation. The stages of the loop closely mirror the process of scientific reasoning engaged by scientists. However, a similar set of inductive processes are at play in many real-world situations (e.g., working an unfamiliar ATM machine, reading a complex nutrition label). Aspects of the loop are directly related to Bayesian models of learning and information gathering (Bonawitz & Griffiths, 2010; Gureckis & Markant, 2009).

elements are properly functioning and aligned the bike moves forward. However, misalignment of even one component can be catastrophic.

Understanding the integrated nature of these cognitive processes is important not just for our scientific understanding of the development of the human mind, but also because of broader educational implications. For example, many educational philosophies emphasize relatively unstructured, self-guided learning environments (Bruner, 1961; Kolb, 1984; Steffe & Gale, 1995). Understanding limitations in children's active inquiry abilities and how each component of such abilities evolves across age can be used to design more effective learning environments for children of various ages. For example, evidence that younger children benefit from assistance in updating their beliefs in response to new evidence would suggest that learning environments for younger children need to provide support for this component of their learning.

The present study attempts to decompose the component processes involved in active inquiry, specifically focusing on the role of belief updating. We tasked five- to ten-year old children to identify a hidden insect in a simple iPad variant of the classic "Guess Who?" game. Children sequentially asked questions to try to identify the hidden target and received truthful answers. Based on prior work reviewed below (e.g., Mosher & Hornsby, 1966), we expected younger children to have difficulty formulating informative queries and thus sought to explore what types of automated assistance might aid children's reasoning strategies. Specifically, we manipulated whether the computer program helped children to use the new evidence that resulted from their queries to narrow down the hypothesis space, or whether they had to reconcile the revealed evidence and the hypothesis space on their own. Our expectation was that helping children to update their beliefs accurately following the receipt of new information would free up cognitive resources and lead to higher quality question-asking. Interestingly, our results opposed this initial hypothesis in that elements which ostensibly made our task more difficult actually improved the quality of children's inquiry behavior and suggest an important refinement of the information processing model summarized in Fig. 1.

1.1. Developmental change in the ability to ask revealing questions

Active inquiry fundamentally depends on the ability of learners to construct actions or queries which gain information (e.g., asking a question of a knowledgeable adult). A now classic way to study this behavior is through experimental tasks based on the 20-questions or 'Guess Who?" game. In the game, the asker (participant) tries to determine a hidden object known only to the answerer (experimenter) by asking a series of yes-or-no questions. Mosher and Hornsby (1966) identified two broad question types commonly used in the game: *hypothesis-scanning* questions test a single hypothesis or specific instance (e.g., "Is it a monkey?"), whereas *constraint-seeking* questions attempt to constrain the hypothesis space faster by querying features that are present or absent in multiple objects (e.g., "Is it soft?"), but that do not directly identify the answer except by virtue of elimination.

A classic finding in this literature is that younger children (e.g., aged 6) tend to ask more hypothesis-scanning questions, while older children (e.g., aged 11) use more constraint-seeking questions, and also tend to find the answer after fewer questions (Mosher & Hornsby, 1966). One explanation is that only older children have developed the ability to focus on the high-level features that group the hypotheses, whereas younger children focus on individual stimuli. Consistent with this viewpoint, manipulations that help children focus on these higher-level features, such as cuing them with basic level category labels instead of exemplar names (Ruggeri & Feufel, 2015), increase the likelihood that young children will generate constraint-seeking questions (see also Herwig, 1982). Further, although young children are often relatively less likely than older children to ask constraint-seeking questions, even younger children (ages 7–9) are more likely to do so when such questions are particularly informative, such as when the hypothesis space is large and there are several equally probable solutions remaining (Ruggeri & Lombrozo, 2014, 2015). These results reinforce the viewpoint described above: having the right set of hypotheses in mind, or being primed with the right level of category information seems to drive more efficient information search.

The behavioral distinction between constraint-seeking and hypothesis-scanning questions can also be studied from the perspective of normative models (Oaksford & Chater, 1994; Nelson, 2005; Tsividis, Gershman, Tenenbaum, & Schulz, 2013). These models attempt to objectively define the "quality" of a question and to see how people's choices compare (see below for a larger discussion). A number of recent studies have explored how children's question asking compared to such models. For example, Nelson, Divjak, Gudmundsdottir, Martignon, and Meder (2014) found that 8–10 year-old children can search a familiar structured domain (people with varying gender, hair color, etc.) fairly efficiently, tending to ask about frequent real-world features that roughly bisected the search space (e.g., gender first). Likewise, Ruggeri, Lombrozo, Griffiths, and Xu (2015) found that children's patterns of search decisions were well-explained in terms of expected information gain (EIG), one popular model from this class which is described below. Perhaps most importantly, these models are highly context sensitive. Rather than arguing that either constraint-seeking or hypothesis-scanning questions are universally "better," these models take into account the current context including the learner's prior belief and the past evidence that has been revealed. This allows much more fine grained predictions. For example, on a given trial a hypothesis-scanning question might be equally informative compared to a constraint-seeking question (e.g., when only two hypotheses remain). In our study we will analyze children's question asking with respect to these models to allow an objective measurement of the quality of their information seeking behavior.