The Role of the Prefrontal Cortex in Inductive Reasoning: An fNIRS Study

Layla Unger (Lunger@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue
Pittsburgh, PA 15213 USA

Jaeh Kim (jaehk@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue
Pittsburgh, PA 15213 USA

Theodore J. Huppert (huppertt@upmc.edu)
University of Pittsburgh, Center for the Neural Basis of Cognition, Clinical Science Translational Institute, Departments of Radiology and Bioengineering
Pittsburgh, PA 15260 USA

Julia Badger (julia.badger@psy.ox.ac.uk)
University of Oxford, Department of Experimental Psychology, 9 South Parks Road
Oxford, UK OX1 3UD

Anna V. Fisher (Fisher49@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue
Pittsburgh, PA 15213 USA

Abstract
This study examined neural activity associated with inductive inference using functional Near Infrared Spectroscopy (fNIRS). Induction is a powerful way of generating new knowledge by generalizing known information to novel items or contexts. Two key bases for identifying targets for induction are perceptual similarity, and rules that specify category-relevant features. Similarity- and rule-based induction have been argued to represent distinct mechanisms, such that only rule-based induction requires executive function processes associated with the prefrontal cortex (PFC), namely: active maintenance of representations and inhibition of salient but irrelevant features. Here, we address the lack of direct empirical evidence supporting this possibility by recording PFC activity using fNIRS while adult participants (n=24) performed an inductive inference task. We found that PFC activity during induction was greater when participants had been taught a category-inclusion rule versus when participants could only rely on overall similarity.

Keywords: inductive inference; fNIRS; PFC

Introduction
Inductive inference is a powerful component of learning because it allows us to use what we already know to derive new information. For instance, knowledge that one’s cat has a four-chambered heart can be generalized to other entities, such as one’s dog. However, in order for inductive inference to provide a useful source of information, targets for generalization must be identified based on bearing some relationship to the known entity.

Behavioral evidence from studies of adult cognition and developmental research suggest a distinction between similarity-based induction, in which targets chosen for inductive inference are those that are globally similar to the known entity, and rule-based induction, in which targets are those that share specific critical features (Sloutsky, Kloos, & Fisher, 2007; Yamauchi & Markman, 2000). This behavioral distinction may emerge because these two forms of inference involve qualitatively different processes. For instance, similarity-based induction may involve a global assessment of the degree to which known entities and potential targets share features, whereas rule-based induction may involve maintenance of the rule-relevant feature in memory and/or inhibition of rule-irrelevant features (Sloutsky, 2010).

One critical implication of this proposal is that rule-based induction should recruit prefrontal cortex (PFC) regions associated with active memory maintenance and inhibition of salient task-irrelevant information (Konishi, Kawazu, et al., 1999; Konishi, Nakajima, et al., 1999) to a greater extent than similarity-based induction. Indirect support for this possibility comes from evidence that rule- and similarity-based processes in other forms of reasoning recruit distinct brain circuitry (Grossman et al., 2002; Koenig et al., 2005; Nomura et al., 2007; Seger & Cincotta, 2002). However, no studies have yet directly tested this possibility.

Here, we first review behavioral evidence for a qualitative distinction between similarity- and rule-based induction, and neuroimaging evidence for distinct patterns of brain activity associated with similarity- and rule-based processes in other forms of reasoning. Then, we present a study investigating whether similarity- and rule-based induction is associated with distinct patterns of neural activity.

Similarity- versus Rule-Based Induction
Studies of both adult and developing cognition have yielded evidence for a qualitative distinction between similarity- and rule-based induction. For instance, in a number of
experiments with adults (e.g., Yamauchi & Markman, 2000) researchers taught participants two artificial bug categories that were each associated with typical anatomical features and a category label, then investigated the basis on which participants either inferred the value of a bug’s occluded anatomical feature in the presence of its category label, or the value of a bug’s label in the presence of all its anatomical features. In the presence of a label, adults tended to make inductive inferences consistent with the label regardless of other anatomical features in a rule-based manner, whereas in the absence of a label, adults made inferences consistent with the degree to which an item’s anatomical features were typical of each category in a similarity-based manner.

Developmentally, several studies suggest that rule- and similarity-based induction emerge at different ages. For instance, recent research has shown that although 4- to 5-year-olds can learn a category-inclusion rule for a set of unfamiliar items, children often fail to use it as the basis for inductive inferences (Badger & Shapiro, 2012; Sloutsky, Fisher, & Kloos, 2015; Sloutsky et al., 2007) (cf. Gelman & Davidson, 2013). With age, children shift towards making rule-based inductive inferences (Badger & Shapiro, 2012). Therefore, whereas similarity-based inference is evident from early childhood, rule-based induction appears to develop gradually.

**Neural Distinction between Similarity- and Rule-Based Reasoning**

As previously noted, the behavioral distinctions between similarity- and rule-based inductive inference may be associated with a neural distinction between the brain circuitry recruited during reasoning. Specifically, it has been suggested that rule-based induction is associated with greater recruitment of processes associated with the PFC than similarity-based induction (Sloutsky, 2010). Indirect evidence for this possibility comes from observations of neural distinctions between rule- and similarity-based processes in other forms of reasoning.

The primary source of indirect evidence for this distinction comes from comparisons of brain activity observed across different studies and tasks. For instance, tasks that require rule learning, such as the Wisconsin Card Sorting Task, yield significant PFC activity (e.g., Konishi, Kawazu, et al., 1999). In contrast, similarity processing is associated with more posterior brain regions (de Beeck, Torfs, & Wagemans, 2008; Weber, Thompson-Schill, Osherson, Haxby, & Parsons, 2009). Similarly, tasks that require implicit extraction of a category prototype from similarities between exemplars involve visual regions that overlap with those associated with similarity judgments (Reber, Stark, & Squire, 1998; Zeithamova, Maddox, & Schnyer, 2008). Together, these findings suggest that rule-based reasoning recruits PFC, whereas processing visual similarity recruits visual cortex regions. However, this contrast between rule- and similarity-based processes is based on a comparison between studies that used very different paradigms.

Indirect evidence from comparisons made within studies comes from a smaller body of research that has primarily focused on novel category learning. These studies have also found neural distinctions between processes that are related to, but do not directly map onto the rule-versus similarity-based induction distinction of interest (Grossman et al., 2002; Koenig et al., 2005; Nomura et al., 2007; Seger & Cincotta, 2002). Many such studies compare rule-based reasoning to processes that do not involve similarity perception, such as learning categories by integrating perceptual information from multiple dimensions. The small subset of studies that have compared rule-vs similarity-based reasoning have done so in domains that require additional processes, such as retrieving previously experienced exemplars or semantic knowledge from memory (Grossman et al., 2002; Koenig et al., 2005). Accordingly, the nature of this distinction varied as a result of the different processes evoked by different tasks.

Such distinctions support the possibility that rule- vs similarity-based induction recruit distinct brain regions. At the same time, the fact that different tasks used across studies yielded different neural distinctions suggests that the nature of a potential neural distinction between similarity- and rule-based induction cannot be inferred from those observed for other forms of reasoning. Therefore, this review underscores the importance of obtaining direct empirical evidence to test the prediction that inductive reasoning, similar to other forms of reasoning, relies on neurally distinct mechanisms associated with rule-based and similarity-based processing.

**Present Experiment**

We focused on differences in PFC activity between similarity- and rule-based induction for two reasons. First, as noted above, the qualitative behavioral distinction between rule- and similarity-based induction may emerge because rule-based induction uniquely requires processes such as focusing on a specific feature to the exclusion of others and maintaining rules in working memory (Badger & Shapiro, 2012; Sloutsky, 2010) that are associated with PFC activity (Konishi, Kawazu, et al., 1999; Konishi, Nakajima, et al., 1999). Second, the most consistent neural distinction observed between rule- and similarity-based processes in other forms of reasoning is that rule-based processing involves PFC activity (Konishi, Kawazu, et al., 1999), whereas similarity processing involves activity in more posterior regions (de Beeck et al., 2008; Weber et al., 2009).

Accordingly, the present experiment tested whether rule-based induction is associated with greater PFC activity than similarity-based induction using functional Near Infrared Spectroscopy (fNIRS), a neuroimaging technology that uses cortical changes in infrared light absorption to measure brain activity. To test this prediction, we recorded PFC activity using fNIRS while adult participants completed one of two versions of an inductive inference task modeled on a paradigm introduced by Sloutsky et al. (2007) and updated with natural kind-like stimuli by Badger and Shapiro (2012). In this paradigm, participants are asked to infer which of two “match” items shares a property attributed to a “target”. In
the “Rule-Given” version, participants were taught the category inclusion rule, whereas in the “No-Rule” version, participants were not taught the rule. We predicted that the Rule-Given version would yield high rates of rule-consistent match inference choices and significant PFC activity, whereas the No-Rule version would yield high rates of similarity-consistent match inference choices and no significant PFC activity.

Method

Participants
Participants were 24 adults (15 female, \(M_{age}\)=19 years) recruited from the undergraduate community at a Northeastern university who received partial course credit.

Materials and Apparatus
Stimuli were presented on a Dell computer screen with physical dimensions 60 cm x 34 cm and pixel dimensions 1920 x 1080. Participants were seated at a desk facing the screen with their heads about 2 feet away from the screen.

Neural activity was recorded using a continuous wave (CW6) real-time fNIRS system (TechEn, Inc.), with 4 light sources, each containing 690-nm (12 mW) and 830-nm (8 mW) laser light, and 8 detectors, to give oxygenation measures in 10 channels on the prefrontal cortex. The sensors were arranged in the layout depicted in Figure 1. Sensors were snapped into a cap strip built from foam sheet and plastic mesh, and connected to the fNIRS system by optic cables. For each participant, the cap strip was positioned on the head, centered on position FpZ by the international 10-20 coordinate system standard, and extending over the Brodmann area 10 (anterior prefrontal cortex) and area 46 (dorsolateral prefrontal cortex) bilaterally. The strip was secured to the head using a neoprene scuba cap.

Induction Task
This task was presented using E-Prime (Psychology Software Tools, 2012). Stimuli were modeled on those created by Badger and Shapiro (2012), and consisted of bugs with five anatomical features that could each take one of two values: Head shape (pointy or round), body shape (triangular or round), body color (purple or green), spot color (brown or grey), and eye color (blue or orange). Of these features, body shape and color took up a larger proportion of the stimuli than the others. Following Badger and Shapiro, one of the smaller features, head shape, was selected as the category rule-inclusion feature. Specifically, bugs with pointy heads were “Rockbugs”, and bugs with round heads were “Sandbugs”. All other features were category-irrelevant.

We used these stimuli to create 16 induction trials, and 51 baseline trials. Each induction trial presented a triad of bugs consisting of a Target, a Rule Match, and a Similarity Match, arranged with the Target on top and the Matches on the bottom to either the right or left (Figure 2).

In half of the induction trials, the Target was a Rockbug, and in the other half, the Target was a Sandbug. The assignment of Rule and Similarity Matches to the bottom right or left locations was randomized separately for triads with Rockbug and Sandbug Targets. In each triad, the Rule Match belonged to the same category as the Target but appeared dissimilar overall, whereas the Similarity match belonged to a different category but appeared similar. To accomplish this, the Rule Match had different values for all features from the Target except head shape and one of the smaller category-irrelevant anatomical features (eye or spot color), whereas the Similarity match had the same values for all features as the Target except head shape and one of the smaller features (see Figure 2). Independently for triads with Rockbugs and Sandbugs as Targets, we randomly assigned whether the small feature shared by the Rule Match and not the Similarity match was eye or spot color. All triads were pseudo-randomized such that no more than two triads with a Target bug from the same category or the Rule and Similarity Matches in the same locations appeared consecutively.

For baseline trials, we used the bugs to create a simple congruent Flanker task (Eriksen & Eriksen, 1974) (see Procedure). Specifically, each baseline trial presented three identical bugs that were all oriented to face either left or right. We approximately equated the number of times all bugs faced either left or right. Baseline trials were integrated with induction trials such that three baseline trials followed each induction trial, and one set of three baseline trials preceded the first induction trial. This ratio was used to ensure that there was a sufficient amount of baseline recording (i.e., approximately 5-10 s per each set of baseline trials, to mirror the length of time on the Induction trials). Baseline trials were pseudo-randomized such that, in each set of three baseline trials, no more than two featured bugs from the same category or bugs facing the same direction.

Figure 1. Probe layout including 4 sources (S1-4) and 8 detectors (D1-8), overlaid on overhead view of a head. Black lines represent source-detector channels. Source-detector distance was 2.8 cm.

Figure 2. Example of induction trial. Top: Target, Bottom Left: Rule Match, Bottom Right: Similarity Match.
The Induction task was presented in two conditions: A “Rule-Given”, condition and a “No-Rule” condition. The No-Rule condition consisted of only the inductive inference and baseline trials. The Rule-Given condition supplemented these trials with two Rule Demonstration slides, and 16 Categorization trials. Each Rule Demonstration slide depicted a pair of either Rockbugs or Sandbugs. Each pair of bugs had opposite values for all non-head shape features. Each Categorization trial presented a single Rockbug or Sandbug, with equal numbers of trials for each category. We assigned half of the Categorization trials for each category to appear before and half to appear after the baseline and induction trials. We pseudo-randomized each subset of Categorization trials such that no more than two bugs from the same category appeared consecutively.

Procedure
Participants were tested in a quiet space. One experimenter administered the Induction Task, and another managed fNIRS data collection (see details below).

Induction Task Participants were randomly assigned to complete either the Rule-Given or No-Rule condition of the task. The procedure for participants assigned to each condition was identical for Inductive Inference and Baseline trials. During Induction trials, participants were told that the Target possessed a novel biological property (e.g., “plaxium hormone”, “tulvex nerve cells”), and asked to decide which of the two Match items shared the property. Baseline trials were modeled on the “congruent” version of the Eriksen Flanker Task, in which participants respond to some characteristic of a central stimulus in the presence of flanking stimuli that possess the same characteristic. This task was chosen as a baseline based on evidence that it elicits relatively little frontal activity (Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002). In our version, participants were asked to point in the direction that the middle bug was facing.

Only Inductive Inference and Baseline trials were presented to participants in the No-Rule condition. In the Rule-Given condition, Inductive Inference and Baseline trials were presented in between an initial Rule Demonstration and Categorization phase, and a final Categorization phase. To demonstrate the rule, the experimenter showed the participant the two Rule Demonstration Slides, and provided the following descriptions of Rockbugs and Sandbugs: “These two pictures are of [Rockbugs]/[Sandbugs]. [Rockbugs live in rocks, and all have pointy heads that they use to dig in the rocks]/[ Sandbugs live in sand, and all have round heads that they use to burrow in the sand]. [Rockbugs]/[Sandbugs] come in many different shapes and colors, but they all have [pointy heads that they use to dig in the rocks]/[ round heads that they use to burrow in the sand].”

To test rule retention, participants were asked to identify the bug on each Categorization trial preceding and following the Induction and Baseline trials.

fNIRS Recording fNIRS data was recorded for each participant using custom data collection software described in Abdelnour and Huppert (2009). The fNIRS cap was first fitted on the head of the participant and the signal quality checked and adjusted if needed to make sure the fNIRS fiber optics made good contact with the scalp of the participant. After initial setup, the fNIRS data was collected at 20Hz at two wavelengths (690nm and 830nm). Following signal quality checking, the experimenter started the induction task. During the induction task, the timing of stimulus onset and offset as presented in Eprime were synced and marked in the fNIRS data by an automated analog signal sent from the computer port (of the stimulus presentation computer) to the fNIRS machine.

Results
Behavioral Results
We first determined that all participants in the Rule-Given condition successfully learned the category inclusion rule (i.e., all Rule-Given participants were 100% accurate on the initial and final Categorization Trials). Responses on the 16 Induction Trials were then analyzed to compare the degree to which participants in the Rule-Given and No-Rule conditions chose the Rule Match. The Rule Match was chosen significantly more often by participants in the Rule-Given condition (M=60.94%) than by than participants in the No-Rule condition (M=10.42%) (t(22)=5.001, p<.0001).

fNIRS Results
Pre-Processing The raw fNIRS data at the two wavelengths were converted into estimates of oxy- and deoxy-hemoglobin using the modified Beer-Lambert law (Cope et al., 1988) with a differential pathlength factor of 6 for both wavelengths. The data was resampled to 4Hz for statistical analysis using an autoregressively pre-whitened weighted least-squares regression model (Barker, Aarabi, & Huppert, 2013). In brief, the stimulus timing of the induction trials are used to construct a hypothesis of the timing of the expected response based on a canonical hemodynamic response. This model is then statistically compared against the data using a general linear model and brain activity is inferred from the statistical tests of the coefficients of this linear model. The iteratively whitened weighted least-squares regression described in Barker et al was used to solve this general linear model and had been previously shown to have substantially improved sensitivity and specificity and control of type-I errors compared to other regression methods for fNIRS data in the presence of physiological noise and potential motion-artifacts from slippage of the head cap. This analysis is similar to the general linear model and statistical parametric modeling methods commonly used in functional magnetic resonance imaging (fMRI) (e.g. SPM; Tak, Uga, Flandin, Dan, and Penny (2016)).

fNIRS Analysis Processed fNIRS data were analyzed to first compare activity during induction trials to baseline trials for
its PFC, a brain region associated with executive functions, to a greater extent than similarity-based induction. Participants completed an inductive inference task in which they could infer that a novel property attributed to a Target item was shared by either a similar looking item from a different rule-based category, or a dissimilar looking item from the same rule-based category. Participants who were taught the category rule prior to the induction task both tended to choose the dissimilar looking same-category item, and revealed significant PFC activity in comparison to baseline. In contrast, participants who were not taught the rule tended to choose the similar looking different-category item, and did not reveal PFC activity above baseline. Finally, participants in the Rule-Given condition showed stronger PFC activity during induction than participants in the No-Rule condition.

These results support the proposal that rule- and similarity-based induction represent qualitatively distinct processes. Specifically, rule-based induction may uniquely require executive functions associated with PFC such as the active maintenance of rules in memory, and/or inhibition of rule-irrelevant input (Badger & Shapiro, 2012; Konishi, Kawazu, et al., 1999; Konishi, Nakajima, et al., 1999; Sloutsky, 2010). This distinction is consistent with similar distinctions observed between rule- vs. similarity-based processes in other forms of reasoning (Grossman et al., 2002; Koenig et al., 2005; Nomura et al., 2007; Seger & Cincotta, 2002). These results also support the proposal that the more protracted development of rule- vs. similarity-based induction implicates a greater role for the slow-maturing PFC in rule-based than in similarity-based induction. The findings presented here provide a foundation for further investigation into currently unresolved questions about inductive inference processes, as described below.

**Limitations and Future Directions**

The present study sets the stage for pursuing several questions that follow on from the present findings and remain currently unresolved. First, the past research that inspired the prediction that PFC activity should be evoked to a greater extent with rule- vs. similarity-based induction also predicts that activity in posterior regions associated with visual processing should be evoked to a greater extent with similarity- vs. rule-based induction. The present study investigated only a one-way dissociation; future research should investigate the predicted double-dissociation to provide further insight into the distinction between rule- and similarity-based inductive inference.

Second, although participants in the Rule-Given condition chose the category match more often than in the No-Rule condition, they did not always do so. This may reflect variability in the degree to which different individuals perform rule-based induction. Future research should therefore test whether such variability both within and across individuals is associated with differences in PFC activity.

Finally, the current work provides a foundation from which to investigate the neural underpinnings of the previously observed distinction between the developmental trajectories of rule- and similarity-based induction. The present study was inspired in part by the possibility that rule-based induction emerges more gradually than similarity-based induction because the former uniquely recruits brain circuitry involving the slow-maturing PFC. However, no research has directly tested this possibility. Because the present study used a child-appropriate paradigm and imaging technology, the approach used here could be used to study the development of the role of the PFC in rule- versus similarity-based induction.

**Discussion**

The purpose of this study was to use fNIRS to investigate the possibility that rule-based induction recruits PFC, a brain region associated with executive functions, to a greater extent than similarity-based induction. Participants completed an inductive inference task in which they could infer that a novel property attributed to a Target item was shared by either a similar looking item from a different rule-based category, or a dissimilar looking item from the same rule-based category. Participants who were taught the category rule prior to the induction task both tended to choose the dissimilar looking same-category item, and revealed significant PFC activity in comparison to baseline. In contrast, participants who were not taught the rule tended to choose the similar looking different-category item, and did not reveal PFC activity above baseline. Finally, participants in the Rule-Given condition showed stronger PFC activity during induction than participants in the No-Rule condition.

These results support the proposal that rule- and similarity-based induction represent qualitatively distinct processes. Specifically, rule-based induction may uniquely require executive functions associated with PFC such as the active maintenance of rules in memory, and/or inhibition of rule-irrelevant input (Badger & Shapiro, 2012; Konishi, Kawazu, et al., 1999; Konishi, Nakajima, et al., 1999; Sloutsky, 2010). This distinction is consistent with similar distinctions observed between rule- vs. similarity-based processes in other forms of reasoning (Grossman et al., 2002; Koenig et al., 2005; Nomura et al., 2007; Seger & Cincotta, 2002). These results also support the proposal that the more protracted development of rule- vs. similarity-based induction implicates a greater role for the slow-maturing PFC in rule-based than in similarity-based induction. The findings presented here provide a foundation for further investigation into currently unresolved questions about inductive inference processes, as described below.

**Limitations and Future Directions**

The present study sets the stage for pursuing several questions that follow on from the present findings and remain currently unresolved. First, the past research that inspired the prediction that PFC activity should be evoked to a greater extent with rule- vs. similarity-based induction also predicts that activity in posterior regions associated with visual processing should be evoked to a greater extent with similarity- vs. rule-based induction. The present study investigated only a one-way dissociation; future research should investigate the predicted double-dissociation to provide further insight into the distinction between rule- and similarity-based inductive inference.

Second, although participants in the Rule-Given condition chose the category match more often than in the No-Rule condition, they did not always do so. This may reflect variability in the degree to which different individuals perform rule-based induction. Future research should therefore test whether such variability both within and across individuals is associated with differences in PFC activity.

Finally, the current work provides a foundation from which to investigate the neural underpinnings of the previously observed distinction between the developmental trajectories of rule- and similarity-based induction. The present study was inspired in part by the possibility that rule-based induction emerges more gradually than similarity-based induction because the former uniquely recruits brain circuitry involving the slow-maturing PFC. However, no research has directly tested this possibility. Because the present study used a child-appropriate paradigm and imaging technology, the approach used here could be used to study the development of the role of the PFC in rule- versus similarity-based induction.
Conclusion
This study investigated whether a proposed qualitative distinction between rule- vs. similarity-based induction (in which the former uniquely involves memory maintenance and/or inhibition) corresponds with a neural distinction in which rule-based induction uniquely recruits PFC. The findings presented here are consistent with the proposed neural distinction, and lay a foundation for further research into the development of rule-based induction.

Acknowledgements
This work was supported by a grant from the Pennsylvania Department of Health's Commonwealth Universal Research Enhancement Program, a Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education, Institute of Education Sciences (R305B040063), and by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition–Scholar Award (220020401) to the last author. We additionally thank Francois Ban and Kevin Long for their help collecting data.

References


