# *Gaze-driven app for infants (DGI18)*

*Jonas Nockert, nockert@kth.se (supervised by Christopher Peters).*

*Intended grade: A*

*Project blog:* `https://lemonad.github.io/ui-for-infants/`

*May 25, 2018*

> [T]here is a clear opportunity to explore the design of apps for very young children, starting as early as 12 to 17 months of age. The evidence presented in this paper suggests a majority [of] children in this age group and older can understand and use basic apps. The research question is how to design them such that they have similar characteristics to beneficial television shows, while helping children build communication, visual, and motor skills, and increase their connections to their caregivers.[1]

[1] Hourcade et al., 2015, p. 1923.

**Contents**



## Background

Hourcade et al., 2015 analyzed videos on YouTube to investigate with what ability infants[2] and toddlers[3] used tablets (see figure 1). It turns out that while most children under 12 months did not seem able to make meaningful use of the devices, the majority of children 12–17 months had reached moderate ability, and by the age of two, 90% had.

[2] Children up to 12 months.

[3] Children one to three years of age.

Childrens' fine motor skills take time to develop. At 12–18 months children begin using their index fingers to point to things, like pictures in books. In that regard, the above results come as no surprise.

Vision, on the other hand, develop earlier. Infants begin tracking moving objects soon after they are born and at about five months, their ability, when it comes to tracking horizontally moving objects, is adult-like[4].

[4] Grönqvist, 2010.

### *Pursuits*

Pursuits[5] is a technique that enables interaction with graphical devices using only gaze. It introduces a new kind of graphical user interface element that is based on movement (see figure 2). A user selects an element by following its specific movements.

[5] See Vidal et al., 2015; and the related video Vidal, 2013.

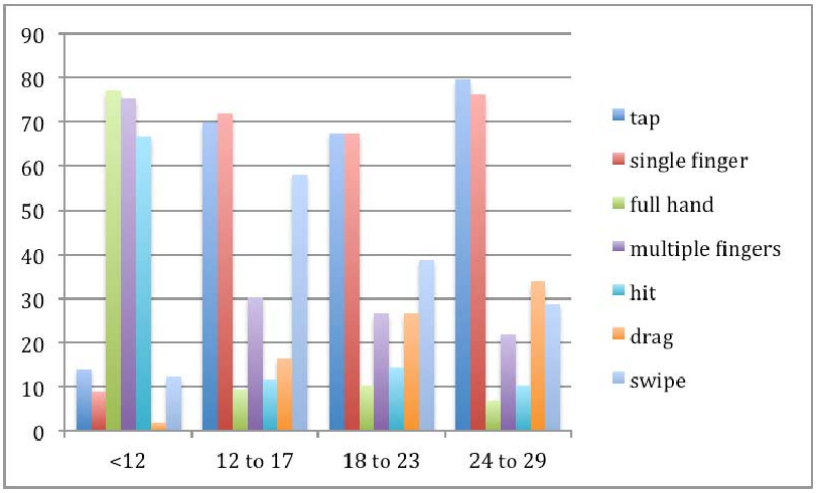


Figure 1: Childrens use of fingers or hands by age group in months (Hourcade et al., 2015, p. 1920).

Pursuits utilizes the smooth pursuit movements of the eye, which is a type of movement that only happens when we are following something with our eyes. Most people can not reproduce this movement on their own, which means that triggering false positives while "just looking" can largely be avoided.

As this technique does not depend on having to identify the position on the screen a user is gazing, only that the gaze is moving in a specific pattern, it seems to be less dependent on exact readings and, better yet, calibration is not necessary as only relative eye movements are relied upon.

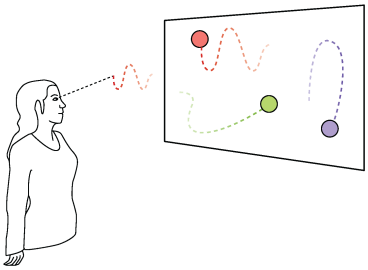

Figure 2: Some examples of Pursuits gaze patterns (Vidal et al., 2015, p. 8).

## Discussion

Suppose that it is the fine motor skills of infants that limit their use of touch-based tablets, not their cognitive abilities. Then, what if an alternative non-touch user interface based on Pursuits was designed, would they be able — and perhaps more importantly — would they want to use that earlier?

The research question is how to design a minimal application that has the potential of verifying it for at least one child. That is, if we can find just *one* child that can learn to use such an interface, it seems reasonable to expect that with additional work, the design could be improved to enable it to work for more children.

The difficulty is that children develop differently compared to each other and that they have good and bad days when it comes to learning something new. In this sense, there is a certain amount of luck involved in order to find the right child at the right time.

To increase the chances of success, we likely need to align the following:

- A consistent and reliable user interface, capable of singling out the random and casual from the intentional.

  Although Pursuits may filter out false positives by using smooth

pursuit eye movements, it utilizes that we, as humans, are drawn to following movement. As it is not possible to ask infants about their intentions, how can one differentiate between them intentionally following an object in order to select it from just following an object because it moved?

Having more than one control seems necessary, and each control matched with a distinct outcome so it could be inferred that the child selects a control over another because it prefers the outcome rather than the control itself. That is, the child needs to be actively using the application rather than just looking at it.

- An application that appeals to parents as something they are interested in seeing their children try out.

  The more children involved on their good days, the better. Parents should preferably be able to use the application with their children on their own, in the comfort of their home.

- Last but definitely not least; infants need to find the application appealing and rewarding enough to use.

**Implementation**

A tablet app seems suitable for the project for a number of reasons. First, the screen size enables large moving objects while still having trajectories resulting in eye-movements that are big enough to be accurately picked up by a front-facing camera. Second, a larger interface should make it easier for an infant to learn to use the interface, compared to, say, the small screen of a mobile phone. Third, tablets, in my experience, represent a type of digital device that parents are most comfortable with their children using. Fourth, developing an app makes it easy to distribute the application to other parents.

Specifically, I will create an app for the iPad 10.5" as it has a front-facing camera with good resolution as well as the processor power to handle computer vision tasks while simultaneously displaying smooth movements

I will use 2D-based SpriteKit for the user interface, Core Image for video processing, Vision for face detection, and DLib for eye detection. Based on previous experience, I will have to write my own pupil tracking algorithm.

The initial goal is creating a basic app like the one sketched in figure 3 where the controls consist of two different horizontally moving images. Statistical data on gaze-following per control will be captured and visualized so it is possible to set up tests to verify reliability on adults as well as tentative tests with toddlers.

Should this prove successful with toddlers, the next step is to extend the testing with infants. One concern is that face and eye detection networks have been trained on pictures of adults and might not work as well with infants due to less distinct facial features.

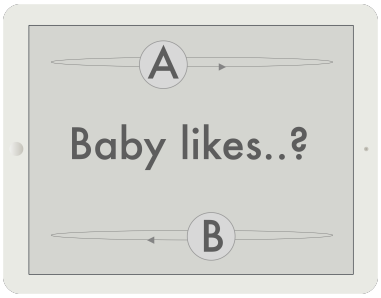


Figure 3: A sketch of the interface of a potential app with two Pursuits controls for infants.

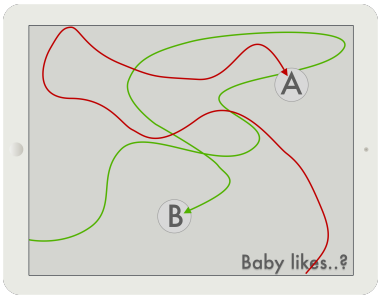


Figure 4: Alternative design with two Pursuits controls for toddlers moving both horizontally and vertically.

Another concern is that they might not be attracted enough by the moving controls or that they get distracted before any meaningful data can be collected. Perhaps some sort of reward mechanism is needed?

Should this succeed, the extended goal is based on the notion that every parent would like to know more about what their child likes and wants. The app could give parents the possibility of uploading their own choices of images, videos or perhaps even music in order to do small and fun experiments of comparing one thing to another. Like to find out if their child prefers apples over carrots, one storybook over another – and jokingly: cats over dogs, C++ over Javascript, or one sports team over another.

If not successful with infants, an alternative goal is to pursue the direction sketched in figure 4: a basic gaze-driven game for toddlers where the premise is to follow more advanced movement patterns, likely using more simultaneous Pursuits controls.

**Evaluation**

All in all, it is unlikely that this project will result in conclusive evidence of infants being capable of learning and willing to use a gaze-driven user interface. It will not give a conclusive negative answer either.

With that said, a user-study of infants and toddlers would obviously give valuable insight, both

- qualitatively: parents' general attitude towards apps for infants and toddlers, their expectations beforehand, reflections on their children using the app as well as their thoughts on directions the app could take in order to enable more interesting experiments.
- quantitatively: do the results differ much between children, how much do the settings and surrounding matter, if the parents use the app on their own, over time, do they get different results?

Together, a user-study would give a clear indication toward the underlying idea being sound or not.

**About me**

I am a third-year Datalogi student at Stockholm University. This is the second time I take on this project and last year, even though I had used iOS and OpenCV for quite some time, as well as read a lot of papers on eye and pupil tracking, I did not have the necessary knowledge to implement a pupil tracking algorithm with the precision and sampling frequency to match minute eye movements.

This year, I come more prepared and will spend the summer working on this project. At my side, I have an enthusiastic 18 month old child who will be more than happy to help out with testing and critique.

**References**

Grönqvist, H. (2010). "Visual motor development in full term and preterm infants." PhD thesis. Uppsala University. ISBN: 978-91-554-7892-6.

Hourcade, J. P. et al. (2015). "Look, My Baby Is Using an iPad! An Analysis of YouTube Videos of Infants and Toddlers Using Tablets." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, pp. 1915–1924. ISBN: 978-1-4503-3145-6.

Vidal, M. (Mar. 2013). *Pursuits: Spontaneous Interaction with Displays*. Youtube. URL: https://www.youtube.com/watch?v=TTVMB59KvGA.

Vidal, M. et al. (Jan. 2015). "Pursuits: Spontaneous Eye-Based Interaction for Dynamic Interfaces." In: *GetMobile: Mobile Comp. and Comm.* 18.4, pp. 8–10. ISSN: 2375-0529.