A

## I

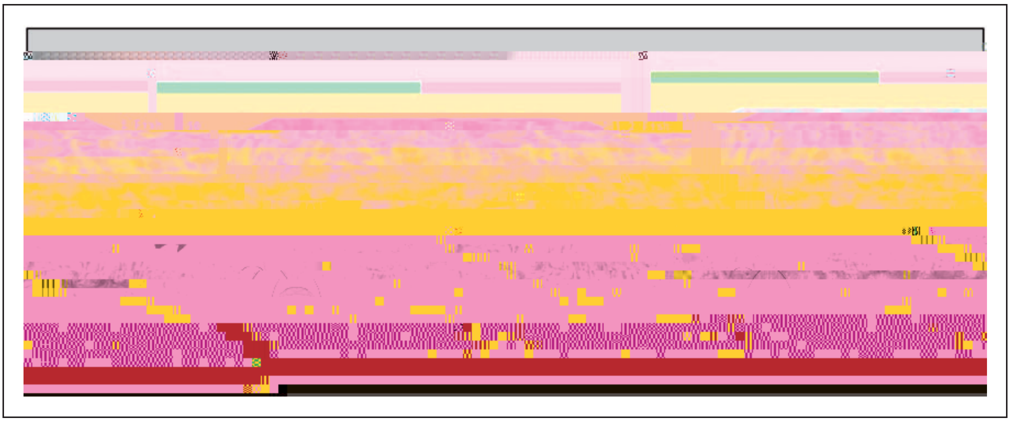
For more than a century, researchers have occasionally collected psychophysical data outside the laboratory, in public or quasi-public settings. Notwithstanding the logistical challenges they present, such environments can be attractive, especially for research focused on developmental trends and individual di erences. Perhaps the best-known example of research done in a public setting is Galton's (1895) study of 9,000 visitors to the London International Health Exhibition. After paying a small fee for the privilege, each visitor's visual acuity, hearing, reaction time, and other functions were measured.

The Living Laboratory, located in Boston's Museum of Science, provides a unique research environment. The Museum invites scientists from local universities and hospitals to use the Laboratory as a venue for engaging visitors in ongoing research that also o ers visitors a valuable educational experience. The Laboratory makes it possible to collect useful data from large and diverse samples of subjects across a wide age range, from school-aged children through senior citizens. For some time, we have been interested in audiovisual integration, particularly the ways in which a correlation between signals' temporal structures promotes their integration (Parise, Spence, & Ernst, 2012). So we took advantage of the Museum environment in order to examine age-related changes in audiovisual integration, and also to test the robustness of observations made previously, in well-controlled laboratory settings (Benussen et al., 2014; Sun, Shinn-Cunningham, Somers, & Sekuler, 2014).

Psychophysical experiments are often time consuming and usually require many repetitions of the same measurement. As a result, test subjects must be available, attentive, and motivated for one or more lengthy testing sessions. Particularly when researchers want to study a wide age range of participants, the repetitive, sometimes monotonous aspects of an experiment can cause participants to withdraw before the needed data have been collected. Withdrawal before testing is complete can render data collected up to that point unusable. To minimize such risks, we embedded our experiment within a simple video game, which was designed to engage and amuse participants while also generating data on interactions between what subjects saw and what they heard (e.g., Abramov et al., 1984; Miranda & Palmer, 2014).

In our video game, ''Fish Police!!,'' players watched as computer-generated fish appeared one at a time and swam rapidly across a virtual river (see Figure 1). As it swam, each fish oscillated sinusoidally in size, at either 6 or 8 Hz. To make the task harder, each fish's path was perturbed by a series of small random vertical displacements. Accompanying each fish was a broadband sound that was amplitude modulated at either 6 or 8 Hz. Subjects were instructed to classify each fish as rapidly as possible using only what they saw, judging whether a fish oscillated at the slower (6 Hz) or faster rate (8 Hz). While making these judgments, subjects were to ignore the concurrent amplitude modulated sound. If, despite these instructions, the sound a ected subjects' responses, we expected that categorizations to be more accurate and faster when visual and auditory signals were Congruent, that is when they shared the same rate, rather than Incongruent, when auditory and visual signals were mismatched in rate.

Once a fish appeared and began its journey across the river, a subject had just 2 seconds to respond before the fish would disappear from view. Then 3 seconds later, a new fish appeared and began its journey across the river. This schedule spawned about 12 fish per minute, which made the game challenging and seemed to promote sustained attention.



equipment needed for research can be set up and left unmolested for repeated use over an extended period. In public settings, that is usually impossible to do. The Living Laboratory's space is time-shared among multiple research projects, each of whom can use the space only a few hours each week. At the end of each day's assigned time, researchers must pack up and remove all the equipment and materials that had been used. To accommodate this requirement, our experiment's game was implemented in Python on an inexpensive touchscreen tablet computer (Samsung Note 10.1) running the Android mobile operating system.

To boost players' enjoyment, Fish Police!! incorporated several features common to video games (Hawkins, Rae, Nesbitt, & Brown, 2013; Miranda & Palmer, 2014). For example, each correct response was followed immediately by a pleasant, rewarding sound (clinking of coins), and each incorrect response brought an unpleasant sound (a short buzz). Additionally, a running total of correct responses, represented by a collection of coins, was displayed at the top of the computer tablet screen (see Figure 1). The length of a green progress bar near the display's top indicated the time remaining before the response deadline. Thanks to these features and the task's inherent challenge, Fish Police!! proved su ciently engaging that of the Museum visitors who began play, only ~10% quit before completing the game. Of these, approximately half were either ushered away by parents or guardians or were interrupted by some other unavoidable event. Moreover, about 25% of the players asked if they could play a second time, a request that we had to decline because there was almost always a line of people waiting their turn to play.

As potential players of Fish Police!! would be unused to the attentional demands of psychophysical experiments, we decided to embed subjects' instructions in a narrative that would be engaging, easily understood, and easily remembered. After being shown the computer tablet on which the game would be played, subjects were told:

> You are going to be the police o cer in charge of a river. One at a time, from either side of the tablet, a fish will appear. They are very nervous, though, because they don't want to be caught by

you! The bad fish will be wiggling fast because they're scared, that's when you tilt the tablet towards you to catch them. If the fish is wiggling more slowly, then it's a good fish because it isn't as nervous, so you can tilt the tablet away from you to let it go. Remember though, since they've been swimming for a while, they're a little bored so they hum –you'll hear this through the headphones I'll be putting on you. Try as best you can to JUST focus on their wiggling to tell if they're a good or bad fish.

Each participant played the game while holding the tablet with one hand on each side of the screen. This made it possible to implement a pair of responses, tilting the tablet either toward or away from the player that had an easily remembered correspondence to the judgment that each direction signaled.

At the Living Laboratory's entrance, a video monitor advertized the game by displaying an image from Fish Police!!'s splash screen. Before commencing play, a player's basic demographic information, initials and age, was entered via the tablet's touchscreen. This information and all data generated during game play were de-identified and uploaded wirelessly by the tablet in real time to a secure server o site. The tablet's built-in accelerometer sampled the tablet's angle of tilt at 60 Hz. The Python script controlling the video game defined a response as a rotation that was 17 contrasted players' performance with Congruent fish to their performance with Incongruent fish, using as metrics the proportion of categorizations that were correct, and the latency of response on correct trials; 60 test participants completed the entire 5-minute game; their ages ranged from 82 years down to 6 years, the youngest age we had permission to test. Figure 2 shows the age distribution of the players.

The 60 fish each player saw and heard were uniformly distributed across four categories defined by the two species of fish (good fish—6 Hz oscillation in size and bad fish—8 Hz oscillation in size), crossed with two types of congruence (Congruent audiovisual signals and Incongruent audiovisual signals). For each subject, this $2 \times 2$ design produced just 15 trials per cell, too few samples for stable estimates of the dependent measures. For example, with only 15 samples and binomial variability, a change of just a single response would produce a swing of more than 13%. To reduce the impact of having so few samples per condition, while responses made to Congruent and Incongruent fish were entered into a one-sample t test, producing $t = 6.89$, which corresponds to $p = 4.128e\text{-}09$ (with $df = 59$) (Figure 3).

We next focused on response times from trials on which correct responses were made.

Each trial's MT was subtracted from its response time, yielding the reaction time for that trial. Thereafter, all chronometric analyses were done on values of this derived reaction time variable. Players' reaction times were significantly shorter for Congruent fish than for Incongruent ones. The mean di erence between the two sets of reaction times was 59.21 ms (95% CI [39.50, 78.95]). A t test on the di erences between the two sets of reaction times produced $t = 6.00$ ($df = 59$), for $p = 1.277e\text{-}07$. So both dependent measures, time and accuracy, produced reliable di erences between players' processing of Congruent and Incongruent fish (Figure 3).

We were interested in the possibility that subjects might have been trading accuracy for speed of response (Beilock, Bertenthal, Hoerger, & Carr, 2008; Heitz, 2014; Wickelgren, 1977). Viewing each fish for a longer time could have allowed additional visual information to accumulate, thereby increasing the proportion of correct responses (Mazurek, Roitman, Ditterich, & Shadlen, 2003; Noppeney, Ostwald, & Werner, 2010). More recently, Teichert, Ferrera, and Grinband (2014) showed that subjects strategically increase response accuracy by delaying the onset of their decisions. Would players whose reaction times were long tend to produce a higher proportion of correct responses? To evaluate this potential connection between players' speed and accuracy, we examined the correlation between the two measures. In Figure 4, the e ect of audiovisual congruence as represented by reaction time is plotted against the e ect of audiovisual congruence as represented by accuracy. These two measures of the congruency e ect were not significantly correlated, $r = -0.006$; the (95% CI included zero [−0.255, 0.244]). So, whatever factors contributed to di erences in the way that players were a ected by audiovisual congruence, these factors did not include adoption of some consistent speed-accuracy strategy.

Previous studies suggested that in humans, audiovisual integration emerges late in the first year of life (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006), but continues to be fine-tuned until somewhat later. Moreover, at least one study (Roudaia, Sekuler, Bennett, & Sekuler, 2013) showed weakened audiovisual interaction in older adults. All of these studies used audiovisual stimuli and tasks that were quite di erent from those embodied in our game. Although we were limited to testing subjects 6 years of age and older, we thought it would be worthwhile to evaluate how audiovisual interaction in Fish Police!! varied with age. Figures 4 and 5 show the results of this inquiry for accuracy and reaction time measures,

## Discussion

Results with each dependent measure were consistent with those from a comparable task studied in a controlled, laboratory setting (Sun et al., 2014). In addition to differences in the level of distractions, the experiments comprising the laboratory study differed in multiple ways from the conditions in the Museum of Science. For example, they differed in size of the display (33° vs. ~14–15° visual angle wide), response modes (button press vs. tablet tilt), number of trials each subject experienced (~300 vs. 60), interfish intervals (~2 vs. 3 seconds), and number of subjects per experiment (10 vs. 60). Table 2 shows that despite these differences, the main outcomes of the present study are not dramatically different from what was seen in the three laboratory experiments. The absence of strong differences among venues and conditions points to the robustness of audiovisual integration that arises from temporally correlated auditory and visual signals.

Moreover, results from the Museum and from the laboratory show that the effect of temporal correlation of auditory and visual signals is strong enough to survive even with no corresponding spatial correlation between visual and auditory information. That is, the spatial information provided by vision (the fish's progress across the virtual stream) was not matched by a comparable change in auditory spatial information (such as a change in interaural time difference).

As explained above, results from the two dependent measures (Figure 4) suggest that differences among players' accuracy could not be explained simply by differences in their strategies for balancing the competing demands for speed and accuracy. It seems likely that the time constraints imposed by the game design (allowing just 2 seconds to respond) created high pressure that might have a particular effect on players tested at the Museum of Science.

interference (Fitts & Deininger, 1954), which was expected to maximize observed differences between responses to Congruent and Incongruent fish.

Planning this project made us mindful of the many differences between a public or quasi-public research environment and the well-controlled dedicated research laboratories in which we have been studying various aspects of audiovisual interactions (Benussen et al., 2014; Sun

## F f

Supported by CELEST, an NSF Science of Learning Center (SBE-0354378). This research was conducted in the Living Laboratory at the Museum of Science, Boston. We thank Justin Harris, manager of the Living Laboratory for his help; Fatima Abu Deeb, Nick Moran and Jonathan Chu provided excellent programming assistance. This study was carried out in accordance with Brandeis University's Committee for the Protection of Human Subjects and the World Medical Association Helsinki Declaration revised in October 2008.

## R

Abramov, I., Hainline, L., Turkel, J., Lemerise, E., Smith, H., Gordon, J., & Petry, S. (1984). Rocket-ship psychophysics. Assessing visual functioning in young children. Investigative Ophthalmology & Visual Science, 25, 1307–1315.

Beilock, S. L., Bertenthal, B. I., Hoerger, M., & Carr, T. H. (2008). When does haste make waste? Speed-accuracy tradeo , skill level, and the tools of the trade. Journal of Experimental Psychology: Applied, 14, 340–352.

Benussen, S., Chow, K., Varghese, L. A., Sekuler, R., Somers, D. C., & Shinn-Cunningham, B. (2014). Bidirectional audiovisual interactions: Evidence from a computerized fishing game. In: Meeting of the Acoustical Society of Americath.StAmeon30-282l7(Laborator29-389.9(af)-305[(th.)8DlaBeor29-391.1(innarect