ADDENDUM: FREE-VIEWING VERSUS DESCRIPTION

After publication of the BubbleView paper, we received questions about whether the description task used with information visualizations was essential to the results we were seeing in Experiment 1 (Sec. 5.1) - specifically, that 90% of free-viewing fixations could already be accounted for by the BubbleView clicks of 10 participants (who were required to provide a description as they clicked around). In this addendum to the paper we address this question:

Motivating question
- How does the task of free-viewing versus description affect the clicks of BubbleView participants?

Stimuli
We sampled 30 images from the set used in Exp. 1.3 (Sec. 5.1). Images were a maximum dimension of 600 pixels to a side and were blurred with a sigma of 40 pixels. We used a bubble radius of 32 pixels.

Method
In a single HIT, participants were shown all 30 images, for 10 seconds each and 2 seconds between consecutive images, with the instructions to "click anywhere you want to look". The resulting clicks collected correspond to a free-viewing task. We collected an average of 60 participants worth of BubbleView click data for each image.

Results
In the description task (from Sec. 5.1), participants made an average of 64 clicks per image over a viewing interval of approximately 3 minutes/image. If we account for the time to write descriptions, participants spent 15-30% of the total task time (of an average of 3 minutes/image) clicking, or approximately 27-54 seconds. That is, 1-2.4 clicks/sec. In comparison, in the free-viewing task, participants made an average of 15 clicks per image over a 10 second viewing interval, or 1.5 clicks/sec. In the original free-viewing task with an eyetracker, participants made an average of 40 fixations per image in 10 seconds of viewing, or 4 fixations/sec, attending to 2-3 times more locations than with clicking.

From Table 1 we see that for the same number of participants, the click data obtained under a description task is significantly more similar to eye fixations (under free-viewing) than the click data obtained under a free-viewing task. With 12 participants, the click maps obtained from the free-viewing task achieve an NSS score of 1.07, extrapolated to increase to 1.19 (95% C.I. [1.18, 1.19]) in the limit (Fig. 1), compared to an NSS of 1.30 already achievable with 12 participants performing a description task (and a limiting NSS score of 1.37, 95% C.I. [1.35, 1.38]). In other words, increasing the number of participants can not compensate for the difference in the quality of clicks generated between the two tasks. Given a description task, participants explore more of the visualization than with the free-viewing task (where participants focused predominantly on titles and main text elements). Note that this result can be confounded with the amount of time participants were given to freely explore each infographic (limited to 10 seconds) in the free-viewing scenario. However, without a task, a longer task may not necessarily keep online crowdworkers motivated.

Take-aways: A description task is recommended over a free-viewing task for BubbleView clicks to better approximate (free-viewing) eye fixations collected using an eye tracker.

Table 1. We evaluated BubbleView clicks at approximating ground-truth eye fixations on 30 images from the MASSVIS dataset, using 2 tasks: description and free-viewing. BubbleView maps were computed with $n = 12$ participants in black, and including all n participants collected for each experiment in gray. The score of the BubbleView maps predicting the ground-truth fixation maps is reported in CC, and the score of the BubbleView maps predicting the discrete fixation locations is reported in NSS. Normalized NSS is calculated by normalizing the NSS score by the inter-observer consistency (IOC) of the eye tracking participants on the 30 images.

Visualizations (ground-truth IOC: 1.42)	CC	NSS	Normalized NSS
Description task	0.87	1.30	92% ($n = 12$)
Free-viewing task	0.72	1.07	75% ($n = 12$)
	0.75	1.13	80% ($n = 57$)

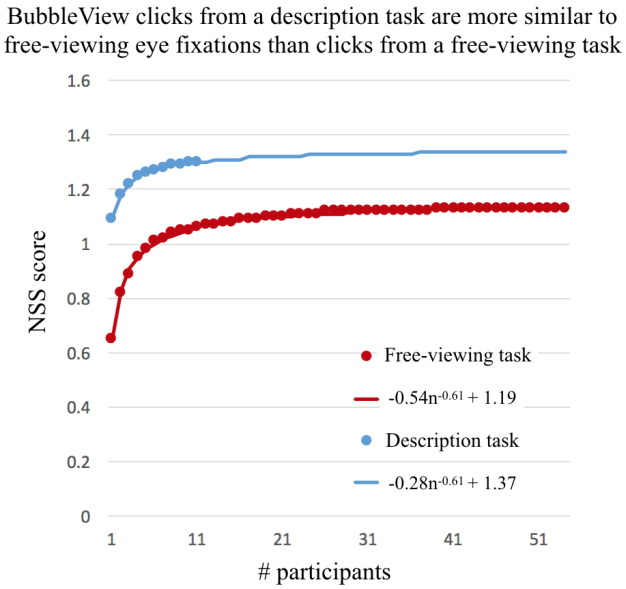

Fig. 1. The NSS score obtained by comparing mouse clicks under free-viewing and description tasks to ground truth eye fixations on a small set of visualizations in the MASSVIS dataset. Each point represents the score obtained at a given number of participants, averaged over 10 random splits of participants and all 30 images used.