

## Abstract

Research suggests that young children typically find the think-aloud method to be difficult (Someren et al., 1994); additionally, it is unclear whether the quality of verbal reporting produced by children also differs across demographic variables, such as SES. This study examines the application of a think-aloud technique in survey-item testing across students in grade 4 (age 9-10). We compared the quality of the think-aloud produced by students of different SES groups (Low and High) by evaluating verbal report length, relevance of the verbal report, and the number of reported problems identified. Results suggest that the quality of information gathered from the think-aloud interviewing method is not affected by SES among grade 4 students.

## Introduction

The need for thorough and comprehensive pre-testing continues to grow in importance for multi-media rich digitally-based assessments (DBAs). Over the last several decades, cognitive interviewing (Tourangeau, et al., 2000) has become one of the most effective methods of eliciting in-depth, multi-faceted feedback from the populations intended as end-users. The think-aloud components of cognitive interviewing requires respondents to verbalize their thoughts as they work through test questions. Think-alouds are designed to uncover a myriad of potential problems that may compromise the quality and validity of responses, including the usability of the DBA, the clarity of instructions, and response options.

However, there is little empirical research evaluating the quality of the think-aloud in children. Past research suggests that the think-aloud method used (concurrent vs retrospective) in grade 4 children may impact the quality of the verbal output obtained. (Sugovic et al., 2017). Similarly, research has shown that a student's SES may affect their ability to accurately recall information (Ginet et al., 2014). In this study, we examined whether a student's SES has a measurable impact on the think-aloud length, relevance, and quality.

## Objective

To evaluate whether students' socio-economic status (SES) affects the quality of the think-aloud produced by grade 4 students.

## Methods

- Twenty grade 4 students participated in cognitive interviews, completing grade-specific science questions.
- Students were provided with consistent and detailed think-aloud instructions, a demonstration, and a think-aloud practice before attempting to answer each question.
- Both High and Low SES students received the same set of six questions consisting of the following:
  - **Subject:** Science NCES Questions administered in 2009
  - **Content Classifications:** 2 Physical Science, 2 Earth and Space Science, 2 Life Science
  - **Question Difficulty:** 2 Easy, 2 Medium, 2 Hard

## Demographics

- A total of twenty grade 4 students participated in cognitive interviews. The sample consisted of a mix of race/ethnicity (Table 1) and gender (Table 2) by SES.
- Low SES classified as  $\leq$  \$50K Household Income
- High SES classified as  $\geq$  50K Household Income

Table 1. Race/Ethnicity by SES

	High SES	Low SES	Total
Black or African American	1	3	4
White or Caucasian	4	2	6
Hispanic or Latino(a)	0	8	8
Asian	2	0	2
<b>Total</b>	<b>7</b>	<b>13</b>	<b>20</b>

Table 2. Gender by SES

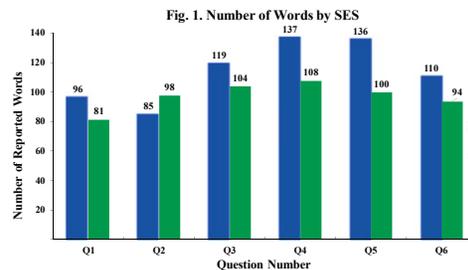
	High SES	Low SES	Total
Female	3	7	10
Male	4	6	10
<b>Total</b>	<b>7</b>	<b>13</b>	<b>20</b>

- Controlled for experience and frequency of technology use (tablet, phone, and computer) across both High and Low SES groups.

## Results

### Number of Words

- There was no significant difference in the verbal report length (number of reported words) between High SES and Low SES students. However, on average, High SES students produced 16.3 more words per question in their verbal reports than Low SES students.

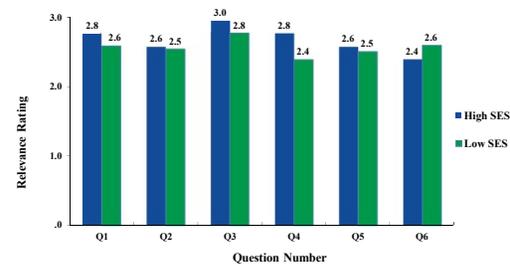


## Results

### Relevance of the Verbal Reports

- No significant difference was found when measuring the relevance of students' verbal think-aloud to the question topic between SES groups. Both High and Low SES students provided an equally relevant verbal think-aloud report to each question.

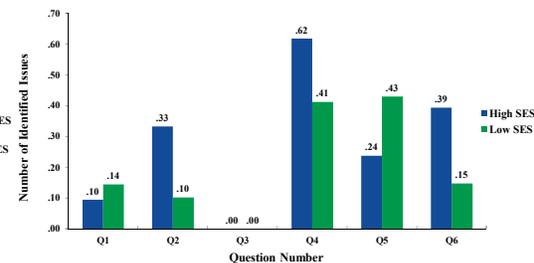
Fig. 2. Relevance of Verbal Reports to Question Topic



### Issues Identified

- There was no significant difference in the average number of reported issues by question between the two groups. Limited findings may be due to low variability in the number of issues identified across both SES groups.

Fig. 3. Average Number of Issues Identified



## Analysis

Students' verbal reports were coded by three coders. The following parameters were used to measure the quality and quantity of each think-aloud:

- **Response Length:** the number of words stated in students' think-aloud.
- **Response Relevance:** overall relevance of reported think-aloud to question topic on a 3-point Likert scale (1 – No, 2 – Somewhat, 3 – Yes).
- **Reported Issues:** average number of issues identified by the student.

1. Results indicate that students' SES does not impact the response length (number of words) of their think-aloud. However, there appears to be a trend indicating that High SES students produced more words over Low SES students.
2. There is no difference in verbal report relevance between High and Low SES students. Students across both SES groups generally provided relevant and high-quality think-alouds.
3. Results suggest that there is no difference in the number of issues identified across the two SES groups of grade 4 students. Results may be affected by the overall low number of issues identified between the two groups across all questions.
4. Preliminary findings suggest that there may be some indication of a difference in verbal report length between Low and High SES students. However, a lack of significant differences may be due to the small sample size as well as an imbalance of students across the two groups. Furthermore, the items tested may pose a limitation given that these are already revised and tested Science NAEP Questions.

Ginet, M., & Brunel, M., & Verkamp, F., & Désert, M., & Colomb, C., & Jund, R. (2014). *Is the Cognitive Interview still efficient with very young children from low SES to testify about a visual event?*. L'Année psychologique, 114, 289-313.

Sugovic, M., Nooraddini, I., Teal, C., Sherech, B. (2017). *Concurrent Versus Retrospective Think-Aloud Method in 4th Grade Children*.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

van Someren, M. W., Barnard, Y. F., and Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London, England: Academic Press.

## More Information:

Please send questions or feedback to Mila Sugovic, [sugovicm@eurekafacts.com](mailto:sugovicm@eurekafacts.com).