Students' PISA test performance and eye movement: A comparison between Germany and Taiwan

Tsai Wen Lin[1*], Stefan Ufer[2], Chao Jung Wu[3]

[1]Department of Educational Psychology and Counseling, National Taiwan Normal University
[2] Mathematics Institute, Ludwig-Maximilians-Universitat Munchen
[3]Department of Educational Psychology and Counseling, National Taiwan Normal University
*skindeep0610@gmail.com

Abstract

Statistical text, used in daily life and even for international assessments, is edited with descriptions of variables, formulae, and graphs that consist of axes, scales, names of variables, and central graphs (Carpenter & Shah, 1998; Elmes, Kantowitz, & Roediger, 2009; OECD, 2013). Statistical text allows readers to integrate the diversity of data information and facilitates decision-making, communication, and persuasion. Further, comprehending statistical text involves integrating and inferring information from verbal and graphical representations simultaneously (Friel, Curcio, & Bright, 2001; Schnotz & Bannert, 2003, 2014; Zhong, Li, Wu, & Lu, 2011). Accordingly, reading comprehension of statistical text can be regarded as a cornerstone for effective learning, and simultaneously for the international comparison of learning effectiveness.

In Taiwan, the statistics curriculum commences with teaching students to collect and categorize statistical data, and then, to read the data involved in graphs as well as generate graphs (Ministry of Education, 2010). In Germany, the statistics curriculum commences with instructing students to understand statistical graphs, particularly reading the data in the graphs (Kultusministerkonferenz, 2004). Therefore, different perspectives have led to differences in the design of the statistics curriculum between Taiwan and Germany. Interpreting statistical data to generate graphs is important at the early stages of the statistics curriculum in Taiwan. Understanding graphs is emphasized at the early stage of the statistical curriculum in Germany. Hence, this

study aims to compare the test performance and reading patterns of 9th grade students between Taiwan and Germany, with an analysis of readers performing problem solving using statistical text.

Participants were 9th grade students in Taiwan and Germany. The research methods comprised an experimental design and semi-structured interviews. Eye movement experiments comprised a pilot study, in addition to a formal experiment. Experiment materials included three PISA items with eight sub-items, including structured and semi-structured ones. Each participant read items randomly and announced their answers when performing problem solving during the experimental procedures. Interviews were held with participants after the experiments.

The results of the test performance as well as eye movements are shown in Table 1.

Table 1   Mean and Standard Deviation of test performance and eye movement

label of items	Test Performance				Total Fixation Duration			
	Germany (n = 25)		Taiwan (n = 21)		Germany (n = 25)		Taiwan (n = 21)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
M43801	.90	(.37)	1.00	(.00)	42.44	(28.32)	37.32	(26.20)
M43802	.90	(.28)	.95	(.15)	47.90	(29.90)	51.89	(34.88)
M91801	.87	(.30)	.83	(.37)	46.18	(28.78)	38.78	(23.19)
M91802	.85	(.37)	.73	.41	38.70	(27.57)	29.12	(19.41)
M91803	.85	(.34)	.68	.44	47.10	(28.14)	36.66	(29.80)
M52501	.77	(.38)	.85	.36	93.70	(30.00)	66.36	(32.34)
[1]**M52502**	**.77**	**(.43)**	**.63**	**.46**	82.39	(38.63)	70.50	(37.09)
[1]**M52503**	**.75**	**(.31)**	**.48**	**.50**	91.89	(34.78)	70.19	(36.88)
Total	6.60	(.81)	6.13	.90	[2]**490.30**	**(83.53)**	**400.31**	**(160.97)**

* [1]The bolded letters are semi-structured items and their test performance, which reveal the statistical significance between countries.

* [2]The bolded numbers are the results of the total fixation duration of the two countries, which reveals the statistical significance between the countries.

First, there was no significant difference between the test performance of Taiwan and that of Germany; $F(1, 45) = 2.46$, $p = .08$. Nevertheless, German students performed better on semi-structured items than Taiwanese students; $F(1, 45) = 4.86$, $p = .03$; $F(1, 45) = 4.24$, $p = .04$. Second, the total fixation duration between the two countries was significantly different, simultaneously on semi-structured items; $F(1, 45) = 3.16$, $p = 05$. Finally, differences in reading pattern were discovered by observing authors during experiments. German students practiced serial reading from graphs to items while Taiwanese students were reading items before graphs and their method of reading the graphs was to browse.

From these results, it could be postulated that German students spent more time reading the text to first understand items with reading graphs completely, before proceeding to the items, which revealed that the participants were inclined to acquire the entire information from graphs and to map the obtained information to items when understanding the items. This reflected the effectiveness of German students' problem solving, particularly in the performance of semi-structured items, with the more difficult items compared to others. In contrast, Taiwanese students had saved time in understanding items by reading the items briefly and browsing the graphs, but this disadvantaged the participants in the interpretation of information in semi-structured items.