



## Effects of Grade Level and Object Size on Students' Measurement Estimation Performance

Hsin-Mei E. Huang<sup>1\*</sup>

<sup>1</sup> University of Taipei, Taipei, Taiwan, REPUBLIC OF CHINA

Received 12 November 2020 ▪ Accepted 2 December 2020

### Abstract

The current study examined the effects of grade level and object size on the ability to estimate the measurement of objects. Fifth- ( $n = 198$ ) and sixth-grade ( $n = 208$ ) students and freshman undergraduate students ( $n = 71$ ) from Taipei city, Taiwan, participated in the study. The findings indicated a tendency for measurement estimation ability to increase with grade level. Overall, estimation performance was most accurate with medium-sized to-be-estimated objects (TBEOs), while estimates of small-sized TBEOs were more accurate than those of large-sized TBEOs. A tendency to underestimate the measurements of large-sized TBEOs was observed in all groups. Those with good estimation abilities showed a preference for using body parts and convenient objects as references. Moreover, the integration of measurement units constructed from previous experience and eyeballing was an essential skill used by good estimators. Suggestions for measurement estimation instruction are discussed.

**Keywords:** estimation strategy, grade, measurement estimation, object size

### INTRODUCTION

Measurement estimation is an important domain of mathematics education (Ministry of Education [in Taiwan], 2010, 2018; National Council of Teachers of Mathematics [NCTM], 2006). Skills in measurement estimation aid students in speedily judging the reasonableness of an estimated answer obtained by physical measurements or in efficiently completing tasks involving measurements (Gooya, Khosroshahi, & Teppo, 2011). Nevertheless, students tend to perform poorly on assessments of measurement estimations, with close to or fewer than half of participants achieving a criterion level of reasonable accuracy (e.g., item 2013-4M3 #10 M1461E1, National Center for Educational Statistics, 2013).

Researchers have argued that grade (level of schooling) (Tretter, Jones, Andre, Negishi, & Minogue, 2006) and object size (Tretter, Jones, & Minogue, 2006; Jones, Tretter, Taylor, & Oppewal, 2008) are influential factors affecting student performance in measurement estimation. Some researchers (e.g., Swan & Jones, 1980; Tretter et al., 2006) have found a positive association between grade and estimation performance. However,

this association has not been consistently observed in some studies on elementary school students (e.g., Desli & Giakoumi, 2017; Montague & Van Garderen, 2003; Pike & Forrester, 1997). One might debate that such inconsistency could result from some specific reasons such as the curricula (Montague & Van Garderen, 2003) and instruction (Desli & Giakoumi, 2017; Pike & Forrester, 1997) provided in school mathematics. Since measurement estimation is included in mathematics curricula in many countries (e.g., Ministry of Education, 2010, 2018; NCTM, 2006; Ruwisch & Huang, 2018), it is worthwhile to carry out a deep inspection of the influence of grade on students' measurement estimation performance.

There is some evidence from studies in the science field that object size influences estimation accuracy. For example, students from the fifth to 12th grades were found to accurately estimate the size of objects between 10 cm and 10 m, but their accuracy declined for both bigger (e.g., 1 billion meters) and smaller (e.g., 1 micrometer) sizes of objects (Tretter, Jones, & Minogue, 2006). Such results are also evidenced in Jones, Tretter, Taylor, and Oppewal's (2008) data collected from novice and experienced teachers. Nevertheless, for estimating the measurements of to-be-estimated objects (TBEOs) of

### Contribution to the literature

- The current study examined effects of grade level and object size on students' measurement estimation ability via mixed methods. How object size impacts students' estimation performance and the strategies used by students who were skilled in estimation across the fifth- and sixth-grade groups and undergraduates were inspected.
- Students' estimation ability grew with the increase in grade level. Overall, students performed the best on estimating medium-sized objects. They were better able to estimate small-sized than large-sized objects. All grade groups tended to underestimate the measurements of large-sized TBEOs. Good estimators showed a preference for using body parts and convenient objects as references for making estimations.

different sizes, whether differences exist in estimation performance between elementary school students and undergraduate students who are at a higher education level remains unclear. Specifically, the sizes of TBEOs that are likely to be regarded either as mathematics-based or as classroom objects (1 cm through 1 m).

Students often over- and under-estimate object measurements (Jones, Forrester, Gardner, Andre, & Taylor, 2012), and their over- and under-estimations may differ depending on the measurement attributes. For instance, young elementary school students (Forrester, Latham, & Shire, 1990) and middle school students (Jones et al., 2012) were found to underestimate length measurements but overestimate area and volume measurements. In contrast, Forrester and Shire (1994) found that underestimations of volume by students increase as the volume of the TBEO increases. Despite the inconsistency of the findings, understanding of students' errors may help diagnose significant ways of mathematical thinking and strategies (Ryan & Williams, 2007). Seeing that errors occur frequently when estimating the measurements of both large-sized and small-sized objects (Tretter, Jones, & Minogue, 2006; Jones et al., 2008), examining whether differences exist in student error patterns when they estimate the measurements of TBEOs of different sizes is essential for gaining insights into students' difficulties in estimating measurements.

Previous studies (Joram, Subrahmanyam, & Gelman, 1998; Montague & van Garderen, 2003) have shown that students who are skilled in measurement estimations tend to use effective strategies (e.g., benchmark comparison) for reaching reasonable estimates. However, types of strategies employed by students who are skilled in measurement estimation for estimating the measurements of school objects of various sizes have not been thoroughly explored. In addition, information on good estimators' thinking about their uses of estimation strategies is lacking.

The current study aimed to explore the effects of grade and object size on student performance in the measurement estimations of classroom objects. Furthermore, student error patterns were inspected for estimating large-sized and small-sized TBEOs, including

length, area, and volume estimations. Another purpose of the study was to investigate strategies used by students who perform well in measurement estimation and to understand successful mathematical thinking while using these strategies.

## THEORETICAL FRAMEWORK

### Mathematical Thinking Involved in Measurement Estimation

Measurement estimation is a mental process of determining the measurement of an attribute of an object without tools (Joram et al., 1998). The idea of measurement estimation is closely related to conceptual understanding of physical measurement because visualization, mental comparisons, and operations of units are common requirements (Joram et al., 1998; Joram, Gabriele, Bertheau, Gelman, & Subrahmanyam, 2005). Thus, both perceptual and inferential abilities are required for the successful estimation of measurements (Forrester et al., 1990). Perceptual abilities include an accurate visualization of the magnitude of objects, while inferential abilities encompass knowledge of the relative magnitude of numbers and measurement units, which pertains to benchmark knowledge (e.g., the use of measure unit), and computational operations. These skills include proportional reasoning, which requires an understanding of multiplicative relationships, and spatial abilities such as the construction of spatial imagery for processing mental representations (Jones, Taylor, & Broadwell, 2009; Jones et al., 2012).

The accurate estimation of length, area, and volume measurements of a three-dimensional (3-D) object is crucial for understanding the physical world. The capability of estimating these spatial measurements facilitates the development of a sense of space occupied by objects (Jones et al., 2012). Furthermore, the measurement of length, area, and volume varies in complexity, and therefore the processes of estimating these attributes differ (Jones et al., 2012). For example, to estimate a length, students across grades tend to imagine iterating a unit of length with an object and then report the estimated length (Lehrer, 2003). Compared to the estimation of length, the estimation of area (or volume)

of a two-dimensional (2-D) (or 3-D) object requires a higher level of proportional reasoning (Jones & Taylor, 2009; Jones et al., 2012) and spatial thinking. Thus, students tend to be better skilled in the estimation of length measurement than of area (Pike & Forrester, 1997) or volume (Huang, 2016).

### Effect of Grade on Measurement Estimations

The role of grade in students' measurement estimation performance has attracted considerable attention. Pike and Forrester (1997) examined 62 students aged 6-11 by means of two sets of tasks through a laptop computer screen, including a textbook format and a story context corresponding to the textbook format. The results revealed no significant grade effects on estimation performance in either length or area estimation. Despite the results, all the students across grades performed better on estimating the textbook-format tasks, which were close to the classroom-based activity, than on the other set of tasks. This suggests that students' prior classroom-based experience of measurement estimations in the textbook context may have had some influence on their estimation performance.

In another study, Montague and Van Garderen (2003) found that the fourth graders, who received the mathematics curricula reflecting the NCTM standards, performed as well as the sixth graders, who did not receive the same mathematics curriculum. Furthermore, in a recent investigation on Greek students' ability to estimate lengths, Desli and Giakoumi (2017) found no significant grade differences in the performance of the third and fifth graders when metric units were used. Although the third and fifth graders were introduced only to metric units in the Greek mathematics curriculum, grade effects were only observed when non-standard units were used. Desli and Giakoumi argued that an inadequate instruction of metric units may constrain the progress of students' competence in measurement estimation.

In contrast, Tretter, Jones, Andre, Negishi, and Minogue (2006) described the importance of learning experience for constructing an understanding of object size from extremely small to very large based on the data of students' performance in estimating scientific phenomena. The results revealed that students from elementary school through doctoral programs were found to universally hold distinct conceptual categories of scale and to use different unit sizes as references for objects of various sizes. For instance, elementary school students used "centimeter" as a measurement unit for small-sized objects (e.g., textbooks) and "meter" (or body length) as a unit for large-sized objects (e.g., an elephant). Tretter et al. also claimed that students' thinking about size demands an understanding of concepts of scale. Greater knowledge of measurements and experiences in direct and indirect comparisons may

help facilitate the development of scale concepts, which in turn improves estimation abilities.

Jones and Taylor (2009) and Tretter, Jones, and Minogue (2006) also suggest that students' estimation ability can be developed through measurement knowledge and physical measurement skills learned in elementary mathematics, including concepts of measurement units, comparing objects and sizes, unit conversion, and the use of estimation strategies. Along with the increase in learning experience, Tretter et al. stated that the ability to use proportional reasoning and visual-spatial skills improves gradually with the enrichment of physical experience of visualizing scale and experience of the mental operation of measurement units. This in turn helps students make progress in conceptualizing the size of objects and scale (Jones et al., 2008).

### Influence of Object Size on Measurement Estimation

Forrester, Latham, and Shire (1990) examined the influences of object size on elementary school students by using various estimation tasks in which the lengths of the measure units and the lengths of the TBEOs used in the area and volume tasks were both less than 30 cm. The findings revealed that object size influenced the performance of young students aged 5-8 ( $N = 70$ ) on the estimation of area and volume tasks, but not on length estimation tasks involving the lengths of steps, jumping, and lying down. In the science field, Tretter, Jones, and Minogue (2006) surveyed the conceptions of spatial scale (distance) of 215 participants ranging in grade from the fifth to 12th grades as well as doctoral students. The findings revealed that across grades students used one or more unit(s) as quick mental reference(s) but they differed in their ability to estimate measurements of objects of different sizes. When the sizes of the TBEOs were close to the range of human scale such as body length, which is a convenient unit for making comparisons, the estimation tasks were accurate. However, the estimation accuracy decreased when the TBEO size was either very small or very large.

Moreover, Huang, Heinze, Ruwisch, Hoth, and Chang (2019) examined 240 seventh to ninth graders' ability to estimate lengths by using a length estimation assessment in which the sizes of the TBEOs were between 1 mm and 1 m. The results revealed that all the students performed better in the context where the TBEOs were not small ( $> 12$  cm) but were touchable than in the small-object contexts ( $\leq 12$  cm) with touchable and untouchable objects. Taken together, object size seems to be an influential factor on students' measurement estimation performance. Despite discrepancies in the definitions of object sizes between studies, the length of objects provided in elementary mathematics textbooks and tasks was commonly between 1 mm and 100 cm (Desli & Giakoumi, 2017; Forrester & Shire, 1994; Huang et al., 2019).

As to the effect of object size on measurement estimation, some researchers (Joram et al., 1998; Tretter, Jones, & Minogue, 2006) have described the different mathematical skills needed for processing the estimations of TBEOs of various sizes, and suggested that the complexity of processing skills may influence estimation accuracy. Joram, Subrahmanyam, and Gelman (1998) suggested that the use of units and unit iteration is the most frequent process for estimating the measurements of an object. Thus, the size of a TBEO may affect the accuracy of an estimate which involves calculations of the number of measure units used. For estimating the measurement of a large-sized object, for example, multiplicative thinking and complicated computations involving more measurement units are needed, and therefore, bigger estimation errors are likely to be made. In contrast, Tretter, Jones, and Minogue (2006) indicated that for estimating the measurement of a small-sized object with the naked eye, being able to accurately compare a measure unit with a TBEO and an understanding of fractions (e.g., a fraction of a meter) are demanded for estimation accuracy. Thus, in addition to unit comparison, the complexity of mathematical skills (e.g., understanding of fractions, proportional reasoning, and calculation) required for estimating the measurements of objects may vary depending on the size of the TBEOs (Jones & Taylor, 2009; Tretter, Jones, & Minogue, 2006). The more complex the mathematical skills needed for processing, the greater the errors that are likely to be made.

### Over- and Under-estimates of Measurements

Estimation error patterns (underestimations and overestimations) may vary with the measurement attributes (length, area, and volume) that are altered. Forrester et al. (1990) found that young elementary school students were prone to underestimate length, while students who had learned multiplication tended to use multiplication and overestimate both area and volume. This tendency has been replicated in the findings of Jones et al.'s (2012) study on 39 middle school students' performance of solving length estimation tasks (e.g., dowel rods and a line drawn on paper).

The findings of Forrester and Shire (1994) exhibited that when a single dimension increased in size, the estimates remained mostly correct based on the performance of 67 elementary school students. However, an increase in two or three dimensions resulted in an increasing tendency to underestimate. This tendency was particularly strong in younger students (aged 8-9,  $n = 24$ ). In contrast, the older students (aged 10-11 years,  $n = 43$ ) were able to compensate for more than one dimension increase. On the one hand, these findings suggest that grade may play a role in estimation error. On the other hand, object sizes may affect whether students make under- or over-estimations.

Jones et al. (2008) investigated 16 experienced science teachers' and 50 novice teachers' concepts of spatial scale. The results displayed that teachers more frequently overestimated the size of objects on the small scale than on the large scale, while tending to underestimate the sizes of objects on the large scale compared to the small scale. Jones et al. argued that estimation errors may result from inadequate knowledge of scale for different levels and less experience of various size scales.

### Strategies for Measurement Estimation

The use of effective strategies helps reach a reasonable estimation accuracy (Jones et al., 2012; Joram et al., 2005). Hildreth (1983) classified the strategies used in the estimation of length and area measurements by 72 students ranging in grade from fifth grade to college freshmen. Strategies were classified as appropriate or inappropriate. The appropriate strategies, which led to more accurate estimations, included the use of benchmarks, prior information about the TBEOs or measurement units, and area formulas. Inappropriate approaches, which resulted in poor estimations, included the use of unsuitable measurement units, guessing, and incorrect procedures for estimating area measures.

Likewise, Montague and van Garderen (2003) reported that estimation performance was associated with the level of sophistication of estimation strategy. Good estimators tended to use sophisticated strategies, which were similar to the type of appropriate strategies categorized by Hildreth (1983), for example, using benchmarks. In contrast, the less skilled estimators preferred to use less sophisticated strategies, which were similar to the inappropriate strategies in Hildreth's (1983) study, such as using unsuitable units and guessing.

In addition, the use of eyeballing for making measurement estimations, which means the use of visualizing only, is also a strategy employed by elementary school students. For example, "I use my eyes" and "I just looked at it and knew." Visualizing, which is needed for students with normal sight while making estimations, is a perceptual-based approach rather than using benchmarks or a justified strategy (Desli & Giakoumi, 2017). Thus, the use of eyeballing is labeled as a less sophisticated strategy based on Montague and van Garderen (2003). Seeing that the use of appropriate strategies is associated with estimation performance, to investigate strategies used by good estimators and to explore ideas that good estimators hold for making estimations when using eyeballing may aid researchers' understanding of how effective thinkers organize measurement knowledge and use strategies related to expertise.

## Research Questions and Hypotheses

The study addressed the following four questions:

1. How do grade and object size influence student performance on measurement estimation?
2. For each grade group, are there differences in the frequency of overestimations and underestimations of the measurement of large-sized objects?
3. For each grade group, are there differences in the frequency of overestimations and underestimations of the measurement of small-sized objects?
4. What strategies are used by good estimators across grades for estimating the measurements of objects?

In the present study, for each grade group, the estimation errors were inspected through students' estimated answers to the length and area tasks involving large-sized objects, respectively. Similarly, the estimation errors were examined through students' answers to the length, area, and volume tasks involving small-sized objects, respectively. Three hypotheses were tested. The first postulates that there is an interaction between grade and object size in measurement estimation performance. The second states that there are differences in the overestimation and underestimation patterns in the specific measurement estimations of the large-sized objects in each grade group. The third states that there are differences in the overestimation and underestimation patterns in the specific measurement estimations of the small-sized objects in each grade group.

## METHOD

### Participants

A total of 477 students from three grade groups participated in the study: the fifth graders ( $n = 198$ , 101 boys and 97 girls) with a mean age of 11.08 years ( $M = 133$  months,  $SD = 3.40$ ); the sixth graders ( $n = 208$ , 98 boys and 110 girls) with a mean age of 12.18 years ( $M = 146.20$  months,  $SD = 5.27$ ), and the first-year undergraduate students ( $n = 71$ , 15 male and 56 female) with a mean age of 19.31 years ( $M = 231.72$  months,  $SD = 14.61$ ). The elementary school groups were recruited from 11 public elementary schools in Taipei city, Taiwan. The undergraduate students were recruited from two departments of the Education College of a public university in Taipei city.

Due to the coordination between testing schedules and academic calendars, the data from the undergraduate groups were collected approximately two months earlier than those collected from the elementary school group. After general testing, one-on-one interviews were conducted with 26 students who were identified as good estimators, including 14 fifth

graders, nine sixth graders, and three undergraduate students. The process for identifying good estimators is described in the following section.

### Instrument

A 17-item paper-and-pencil assessment developed by Huang (2016) was used to collect data. The assessment included six length estimations, six area estimations, and five volume estimations. These items were fill-in-the-blank, drawing a line according to a prescribed length, and multiple-choice, which required judgements on the reasonableness of answers (see [Appendix A](#)).

The estimation assessment included three sections categorized by the size of the TBEOs. (1) The small-sized (S) section. The S section contained six items in which the TBEOs had lengths (for length tasks) or side lengths (i.e., the length of one dimension of the 2- or 3-D TBEOs) between 1 and 10 cm. The section included one length item, two area items, and three volume items. (2) The medium-sized (M) section. The M section consisted of five items in which the TBEOs had lengths or side lengths between 11 and 50 cm. This section included three length items and two volume items. (3) The large-sized (L) section. The L section contained six items in which the TBEOs had lengths or side lengths between 51 and 100 cm. The L section included two length items and four area items.

According to Huang's (2016) study, the estimation assessment instrument was examined by a panel of mathematics researchers and educators, including five experts, for validity (content). The split-half reliability of the assessment was 0.74 when tested with a sample of students in the fifth- and sixth grades ( $N = 201$ ). The average values of task difficulty of the S, M, and L sections were 0.45, 0.51, and 0.38, respectively.

The estimation assessment was presented on an A3-sized worksheet and could be completed in approximately 40 minutes. The characteristics of the estimation items, TBEOs, size sections, and answers for the estimation assessment are presented in [Appendix B](#).

The TBEOs (or benchmarks prescribed) were presented in three ways as follows: (1) Pictorial presentation. The TBEOs were presented as pictures with (Q15) or without (Q1, Q2, Q7) a physical demonstration of the benchmarks prescribed. (2) Physical presentation. The objects were presented in two ways: (a) both the TBEOs and benchmarks prescribed (Q5-1, Q5-2, and Q11) or (b) the TBEOs only (Q3, Q4-1, Q6, Q8-1, Q10, Q12-1, and Q14). (3) Other presentations. Neither a TBEO nor a benchmark prescribed was presented in this category. For example, drawing a straight line according to a given length (Q13) or drawing a rectangle (or square) (Q16) that matched a given area measurement, or finding one classroom object with an area of  $1\text{-m}^2$  (Q9). The majority of the TBEOs and some benchmarks prescribed were statically presented

on the front desk or blackboard in the classroom. Students were permitted to touch, but not move the objects under the surveillance of a trained research assistant.

To collect strategies used for estimating the length, area, and volume measurements, there were three open-ended questions that required students to write down the estimation methods used. The three questions included one each for the length (Q12-1) and area (Q8-1) estimations of large-sized objects, and one for the volume estimation of a medium-sized object (Q4-1).

Good estimators (see definition below) were asked a question at a follow-up interview "What methods do you use most frequently for estimating the length or area or volume of an object?" The interviewees' responses were audio-taped and transcribed for analysis.

### Scoring, Identification of Good Estimators, Classifications of Estimations, and Strategies

*Scoring.* Each item was scored from 0 to 2 points depending on various percent errors of estimation with reference to Swan and Jones' (1980) and Coburn's (1987) suggestions. A score of 2 indicated a "reasonable" estimate within  $\pm 10\%$  of the actual value, whereas a score of 1 indicated an "acceptable" estimate between  $\pm 10\%$  and  $\pm 25\%$  of the actual value. An estimate of greater than  $+25\%$  or lower than  $-25\%$  of the actual value was considered "inappropriate" and given a score of 0. All of the items were evaluated by the aforementioned values of percent error, excluding the item (Q9) referring to the classroom object of  $1\text{-m}^2$  in area.

The position of a target item and the distance of the target item from the estimator may affect the accuracy of an estimate (Sternberg & Sternberg, 2017); therefore, the values of percent error of the item (Q9) referring to the classroom object were extended as in Huang's (2016) study as follows: 2 points were given for an estimate within  $\pm 25\%$  of the actual value, while 1 point was given for an estimate between  $\pm 25\%$  and  $\pm 50\%$  of the actual value. If an estimate was greater than  $+50\%$  or lower than  $-50\%$  of the actual value, the item was scored as 0. The total possible score of the entire measurement estimation assessment was 34 points. The sub-total score of the S, M, and L sections were 12, 10, and 12 points, respectively.

*Identification of good estimators.* For each grade group, the participants who scored within the top 25% of the total assessment score were identified as good estimators. The number of good estimators in each group was as follows: the fifth-grade group, 49 students who scored  $\geq 17$  points; the sixth-grade group, 52 students whose scores were  $\geq 20$  points; and the undergraduate group, 19 students whose scores were  $\geq 25$  points.

*Frequency of over- and under-estimation.* The frequencies of over- and under-estimations were

calculated based on students' responses to the following five fill-in-the-blank items. For the length estimation, over- and under-estimations were examined by two items: Q1 (a small-sized TBEO) and Q12-1 (a large-sized TBEO), respectively. For the area estimation, over- and under-estimations were examined by Q7 (a small-sized TBEO) and Q8-1 (a large-sized TBEO), respectively. For volume estimation, over- and under-estimations were examined by Q5-1 involving a small-sized TBEO.

An estimate  $\leq -25\%$  of the actual value was defined as an underestimate, while an estimate  $\geq 25\%$  of the actual value was defined as an overestimate. The reasonable and acceptable answers, which were classified as acceptable answers, and non-numeric responses (e.g., a blank) were excluded from the calculation of over- and under-estimate frequencies.

*Categorization of estimation strategy.* The estimation methods described by the good estimators were classified based on the eight categories of estimation strategies (Hildreth, 1983; Joram et al., 1998). (a) Eyeballing. Using the eyes or looking at objects. (b) Body parts. The use of body parts as references, such as the width of a finger or the length of a hand span. (c) Previous knowledge (or experience). Using a form of judgement based on experience or previously learned measurement knowledge such as the use of formulas for area and volume measurements. (d) Mental ruler. Employing a mental reference unit (e.g., 10 cm or 1 m) for measurement comparison. (e) Object. Using objects that are nearby as measurement reference units. (f) Guessing. Guessing represents a gross estimate without methodological thinking. (g) Mixed methods. Combining two or three of the strategies mentioned above. For example, integrating the use of eyeballing and a body part reference. (h) Others. This category contained a response of "Do not know" or unclear description such as "a unit" or a blank.

### Data Analysis

Repeated measures ANOVAs were performed to examine the effects of grade and object size on estimation. For each grade group,  $\chi^2$  nonparametric tests were conducted to examine differences in frequencies of over- and under-estimations for each measurement estimation task. Lastly, the types of strategies used by the good estimators were coded and the frequencies of various strategies were calculated.

Two trained raters worked independently to score written answers on the assessment for approximately 11% of the participants ( $n = 54$ ). Interrater agreement on the scores of the estimation questions reached 100% agreement. Moreover, a Cohen's Kappa analysis testing the reliability of the coding of written strategies was assessed at 0.91,  $p < 0.001$ .

**Table 1.** The Average Means and Standard Deviations for the Students' Scores by Grade and Object Size

	n	S section		M section		L section		Entire test		F	p	ES
		M	SD	M	SD	M	SD	M	SD			
Grade										43.96	< .001	.16
Fifth-grade group	198	.76	.42	1.07	.36	.72	.38	.83	.02			
Sixth-grade group	208	.91	.48	1.14	.32	.80	.38	.94	.02			
Undergraduates	71	1.23	.44	1.42	.37	1.03	.43	1.22	.04			
Object Size		.90	.48	1.15	.36	.80	.40			118.49	< .001	.20

**Table 2.** The Frequencies of Estimated Answers by Grade across Attributes and Object Sizes

Grade group	Estimated Answer	Length Estimation				Area Estimation				Volume Estimation	
		Large		Small		Large		Small		Small	
		f	%	f	%	f	%	f	%	f	%
Fifth-grade group (n = 198)	AC	101	51%	119	60%	29	15%	74	37%	104	52%
	NR	7	3%	0	0%	6	3%	3	2%	3	2%
	OE	19	10%	43	22%	16	8%	60	30%	63	32%
	UE	71	36%	36	18%	147	74%	61	31%	28	14%
Sixth-grade group (n = 208)	AC	132	63%	130	63%	48	23%	112	54%	124	59%
	NR	4	2%	0	0%	5	3%	2	1%	0	0%
	OE	14	7%	48	23%	18	8%	44	21%	49	24%
	UE	58	28%	30	14%	137	66%	50	24%	35	17%
Under-graduate group (n = 71)	AC	56	79%	60	85%	26	37%	45	63%	61	86%
	NR	0	0%	0	0%	0	0%	0	0%	0	0%
	OE	0	0%	3	4%	2	2%	12	17%	3	4%
	UE	15	21%	8	11%	43	61%	14	20%	7	10%

Note. AC = acceptable; NR = non-numeric response; UE = underestimate; OE = overestimate; f = frequency; % = percent

## RESULTS

### Students' Estimation Performance in Various Size Sections across Grades

Table 1 presents the average means and standard deviations of estimation performance on one item from each size section, and overall performance on the entire estimation assessment by grade. As can be seen in Table 1, performances in all three grade groups followed a similar trend. Performance was greatest in the M section of the assessment, followed by the S and then the L sections.

The effect of grade and object size on estimation performance was examined by a 3 (grades) × 3 (sizes) ANOVA with repeated measures on object size. A Mauchly's test indicated that the assumption of sphericity had been violated,  $\chi^2(2) = 7.34, p < 0.05$ ; therefore, the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ( $\epsilon = 0.99$ ). The results showed statistically significant main effects of grade ( $F[2, 474] = 43.96, p < 0.001, \eta^2 = .16$ ) and object size ( $F[1.99, 941.44] = 118.49, p < 0.001, \eta^2 = 0.20$ ) but there was no grade-by-size interaction, ( $F[3.97, 941.44] = 2.10, p = 0.08$ ). Bonferroni corrected post hoc tests revealed that the undergraduate group outperformed both the fifth- and sixth-grade groups, and the sixth-grade group performed better than the fifth-grade group.

For the effect of object size, Bonferroni corrected post hoc tests revealed that all grade groups performed best

on the M section ( $ps < 0.001$ ) and performed better on the S section than on the L section ( $ps < 0.001$ ). Multiple comparisons showed the following results: (a) the fifth-grade group performed best on the M section, while performances on the S and L sections were not significantly different; (b) both the sixth-grade and undergraduate groups obtained the highest scores on the M section and performed better on the S than on the L section.

### Comparisons of the Frequency of Over- and Under-Estimations

Table 2 shows the frequencies and percentages of the four types of estimated answers (i.e., acceptable, blank, under-estimate, and over-estimate) in each grade by measurement attribute and object size. For each grade group,  $\chi^2$  nonparametric tests were conducted to compare the differences between underestimate and overestimate frequencies across measurement attributes and sizes.

The patterns of estimation error for the large-sized TBEOs are as follows: (1) There was a higher frequency of underestimations than overestimations for the length assessment among the fifth graders ( $\chi^2[1, 90] = 30.04, p < 0.001$ ) and sixth graders ( $\chi^2[1, 72] = 26.89, p < 0.001$ ). There were no overestimates for the length estimation assessment in the undergraduate group and therefore statistical analyses were not performed. (2) There was a higher frequency of underestimations than overestimations for the area estimation assessment

**Table 3.** Strategies Used by the Good Estimators in the Assessment of Length, Area, and Volume

Strategy	Fifth-grade group ( <i>n</i> = 49)					Sixth-grade group ( <i>n</i> = 52)					Undergraduates ( <i>n</i> = 19)				
	LE	AE	VE	Total	%	LE	AE	VE	Total	%	LE	AE	VE	Total	%
Eyeballing	3	4	1	8	6%	10	7	9	26	17%	0	1	2	3	5%
One body part	16	17	20	53	36%	10	7	8	25	16%	4	3	6	13	23%
Previous experience	1	5	5	11	7%	1	3	6	10	7%	0	0	0	0	0%
Mental ruler	4	3	6	13	9%	6	2	4	12	8%	2	1	1	4	7%
Objects	18	12	7	37	25%	18	19	4	41	26%	12	12	7	31	54%
Guessing	1	0	0	1	1%	3	2	2	7	4%	0	0	0	0	0%
Mixed two types	0	7	6	13	9%	2	10	16	28	18%	1	1	2	4	7%
Others	6	1	4	11	7%	2	2	3	7	4%	0	1	1	2	4%
Total	49	49	49	147	100%	52	52	52	156	100%	19	19	19	57	100%

Note. LE = length estimation; AE = area estimation; VE = volume estimation

among all grade groups (fifth graders,  $\chi^2 [1, 163] = 105.28, p < 0.001$ ; sixth graders,  $\chi^2 [1, 155] = 91.36, p < 0.001$ ; undergraduates,  $\chi^2 [1, 45] = 37.36, p < 0.001$ ).

With respect to the pattern of estimation errors for the small-sized TBEOs, the results of each group are as follows. (1) There was a higher frequency of overestimations than underestimations for the length assessment among the sixth graders ( $\chi^2 [1, 78] = 4.15, p < 0.05$ ). However, the frequencies of underestimation and overestimation of the length assessment did not differ among the fifth graders ( $\chi^2 [1, 79] = 0.62, p = 0.43$ ) or the undergraduates ( $\chi^2 [1, 11] = 2.27, p = 0.13$ ). (2) There were no differences in the frequencies of underestimations and overestimations in the area assessment in any group (fifth graders,  $\chi^2 [1, 121] = 0.01, p = 0.93$ ; sixth graders,  $\chi^2 [1, 94] = 0.38, p = 0.54$ ; and undergraduates,  $\chi^2 [1, 26] = 0.15, p = 0.70$ ). (3) There was a higher frequency of overestimations than underestimations among the fifth graders in the volume assessment ( $\chi^2 [1, 91] = 13.46, p < 0.001$ ). However, the frequencies of underestimation and overestimation in the volume assessment did not differ in the sixth-grade group ( $\chi^2 [1, 84] = 2.33, p = 0.13$ ) or in the undergraduate group ( $\chi^2 [1, 10] = 1.60, p = 0.21$ ).

### Strategies Used by Good Estimators

Table 3 shows the frequencies of various strategies used by the good estimators in each grade group. Across grades, about 80% of the good estimators or more used effective strategies that facilitated estimations, including body parts as references, previous experiences, mental rulers, objects, and a mix of two types of strategies. Furthermore, the use of body parts and objects as references was used more frequently than other strategies. The sixth graders used eyeballing and a mix of two types of strategies more frequently than the other groups. Despite this, some good estimators in the elementary group but not in the undergraduate group used "guessing." Furthermore, a few good estimators in all grades reported the use of strategies that fell into the "others" category which included unclear or ineffective methods.

As can be seen in Table 3, the good estimators in the elementary school group used mixed strategies more frequently in the assessments of area and volume estimations than in the assessment of length estimation. In contrast, the frequencies of mixed strategies used by undergraduates for area and length estimations seemed equal. The types of estimation strategies used by the undergraduates were similar to those used by the elementary school groups with the exception that the undergraduates did not use previous experience or guessing.

The follow-up interview data provided further insights into the interviewees' estimation strategies, including how they used measurement units for mental operations when using eyeballing and mental rulers. Ten interviewees who used eyeballing and five interviewees who used mental rulers indicated that they imagined standard lengths (e.g., 1 cm, 10 cm, 15 cm, and 1 m) and used them as measurement units (or mental rulers) in the mental operation of unit iterations. They also expressed that the size of the TBEO affected which mental ruler they would use as a reference, in spite of the frequent use of small units such as 1 cm, 1 cm<sup>2</sup>, and 1 cm<sup>3</sup>. For example, one fifth-grade interviewee who used eyeballing indicated,

"If the object looked very big, I would measure it using a mental image of 1 m. I would first (mentally) measure the length and width before applying a formula."

Similarly, one sixth-grade interviewee who used eyeballing described,

"I usually try to estimate by eyeballing it. I mostly use the images of a 1-m stick and a 10-cm ruler in my head. If I have some objects nearby that are about one meter or something, I would probably use those things to estimate."

Most of the interviewees were familiar with the lengths of their body parts and pencils, and they tended to use these references flexibly for comparison. For

example, one sixth-grade interviewee who used body parts expressed,

“If the object I see is very long or large, I would estimate using the number of footsteps. If it is something smaller, I would compare it to something around one centimeter... or fifty centimeters, which is half a step. Because one centimeter is about the thickness of my finger now. Or, I would use comparison because I knew the size of a 1-cm<sup>3</sup> cube.”

Furthermore, most of the elementary school interviewees knew the size of a 1-m<sup>2</sup> square. This knowledge was based on previous learning experiences from activities such as measuring and making a 1-m<sup>2</sup> space using sheets of newspaper.

Interestingly, elementary school interviewees tended to describe the use of guessing when they were uncertain about the accuracy of an estimated answer or when a mental ruler was used for estimations without confidence, based on the interview data. For example, a fifth-grade interviewee who used the width of her finger as a measure unit without complete certainty of the unit length expressed “roughly... roughly..., guessing” for estimating measurements.

Additionally, the interviewees stated that they estimated the side lengths of the 2- and 3-D TBEOs and used measurement formulas for estimating area and volume. However, some interviewees said that they did not mention the use of measurement formulas as much when they described their methods for estimating lengths in the written assessment.

## DISCUSSION

The current study examined the effects of grade and object size on students’ performance of measurement estimations of school objects. The results indicate that grade and object size individually influence estimations, but that there are no interaction effects between the two factors. Overall, the undergraduate students performed significantly better than the elementary school students, while the sixth graders outperformed the fifth graders. All groups performed best in the medium-sized TBEO tasks. Moreover, both the sixth graders and undergraduates performed better in the small-sized TBEO tasks than in those involving large-sized TBEOs, while the fifth graders performed equivalently in the tasks estimating small and large TBEOs. Thus, the current findings did not support the hypothesis that there are interaction effects between grade and object size on students’ estimation performance.

For all grade groups, the frequency of underestimations was greater than that of overestimations in tasks involving large-sized TBEOs. The results support the second hypothesis that there are differences in the frequencies of overestimations and

underestimations across measurements in each grade group for large objects. However, the frequencies of underestimations and overestimations did not differ for most tasks involving small-sized TBEOs across measurements in each group. There were two exceptions: the sixth-grade group overestimated the length of the item and the fifth-grade group overestimated the volume of the item. Thus, the results partially support the third hypothesis.

Good estimators used similar strategies in their estimations of large-sized objects, such as the use of body parts or objects as references. Guessing was used by some good estimators in the elementary school group. Furthermore, low frequencies of the use of unclear or ineffective strategies were found across grade groups. Most of the interviewees were inclined to use measurement units such as 1 cm, 10 cm, and 1 m for comparing the TBEOs when using the eyeballing or mental ruler strategies. The following discussion centers on specific findings, and instructional implications are suggested to strengthen students’ ability to estimate measurements of school objects.

### The Role of Grade and Object Size in Estimation Performance

The findings indicate that estimation performance was improved with increase in grade. The effects of grade on estimation performance is consistent with previous results (Tretter, Jones, & Minogue, 2006; Tretter et al., 2006). With an increase in grade comes an increase in base knowledge of scales in conjunction with an understanding of proportional reasoning (Jones et al., 2009; Jones et al., 2012) and experience with physical measurement. Interviews with the good estimators in the current study provide evidence that experience with real measurement helps the development of measurement estimation skills.

Despite the differences among grade groups, overall, students performed poorly on the estimation assessment in general, given that most of the TBEOs were touchable. The overall percentage correct was approximately 42% for the fifth graders, 47% for the sixth graders, and 61% for the undergraduates. Bright (1979) suggested that touchable TBEOs or measurement units lead to more reasonable estimates due to a decrease in the necessary mental operations (i.e., imagining or recalling the sizes of TBEOs); therefore, the results suggest that the participants in the current study were below average estimators. The conceptual understanding of scale and size (Jones et al., 2008), along with activities that help build a set of scale references (Tretter et al., 2006) should be highlighted in elementary school mathematics.

All grade groups performed best on estimations of medium-sized TBEOs that had lengths between 11 and 50 cm. According to the scoring scheme used in the current study, all groups were able to estimate within

10% to 25% of the actual measurements of the medium-sized TBEOs, reaching acceptable accuracy. These findings are similar to those obtained by Ruwisch, Heid, and Weiher (2015) who found that fourth-grade students performed better on estimation of medium-sized length measurements (i.e., between 8 and 46 cm).

Furthermore, both the sixth graders and undergraduates in the current study made more accurate estimations of small-sized measurements (i.e., between 1 and 10 cm in length) than large-sized measurements (i.e., between 51 and 100 cm in length). However, this did not hold true for the fifth graders. The possible reasons why the fifth graders performed differently from the other two groups are discussed later. Overall, these findings are in agreement with those of another study (Jones et al., 2009), suggesting that estimating large objects is challenging for students.

In the current study, most of the TBEOs were classroom objects that were familiar to students. The effects of object size on estimation performance may result from operations of unit iteration, which are used frequently for measurement estimations (Desli & Giakoumi, 2017; Joram et al., 1998). Calculations for obtaining estimated measurements of the small-sized TBEOs are less complicated than those needed for estimating the measurement of the large-sized TBEOs used in the study. Therefore, the cognitive demands needed for comparing a measurement unit of small-sized objects can be reduced (Joram et al., 1998), which in turn increases estimation performance. The complicated operations of unit iteration and calculations necessary for measurement estimations may lead to students' difficulty in making accurate measurement estimations of large-sized TBEOs.

With respect to the result that measurement estimation performance was more accurate for medium-sized compared to small-sized TBEOs, estimation difficulty is related to the object attribute (length, area, or volume) to be estimated (Jones et al., 2012). As the M section consisted of more length estimation tasks and fewer area (or volume) estimation tasks than the S and L sections and therefore, the task difficulty level of the M section was lower than that of the other sections as shown in the previous section, this may explain why the results were better for this section. A higher level of proportional reasoning is demanded for estimating measurements of 2- and 3-D objects than of objects of just one dimension (Pike & Forrester, 1997). Thus, area and volume measurement estimations are more difficult for elementary school students than length measurement estimation.

In contrast, the total number of tasks that required estimation of area and volume measurements in the S section (five tasks) and L section (four tasks) was close. It can, therefore, be reasoned that the sixth graders and undergraduates performed better in the S section than in

the L section due to the size discrepancy of the TBEOs (not due to the attribute to be estimated).

The fifth graders, on the other hand, obtained similar scores on both the S and L sections. One possible explanation for this finding is that more than half of the estimations in the S and L sections pertained to area and volume, which may be equally challenging to fifth graders. Skills involved in volume measurement are formally taught at the fifth-grade level in Taiwan (Ministry of Education, 2010). Due to inexperience in estimating volume measurements, the fifth graders had limited, but equivalent, competencies in the S and L sections. Further studies examining the effects of object size and measurement attributes on measurement estimation performance are needed.

### **Frequency of Over- and Under-Estimations by Object Size**

In agreement with Jones et al. (2008), the current study found a high frequency of underestimations of length measurements of large-sized TBEOs. This tendency to underestimate the area measurement of large-sized TBEOs was also found in the pattern of estimation errors made by all three grade groups. Similarly, Forrester et al. (1990) found that for area (and volume) estimations, as the ratio between size of measurement unit and size of TBEO increased, estimates shifted from overestimations to underestimations. An inclination to underestimate the measurements of large-sized objects may possibly result from the use of small measurement units, as revealed in the interviews. Using small measurement units for estimating large objects may easily lead to computational errors, due to the complex processing of unit iterations.

Despite a tendency to underestimate the measurements of large-sized TBEOs, there was no similar evidence for small-sized TBEOs. The findings revealed that the frequencies of under- and over-estimations in all three grade groups, for most of the small-sized objects were close, with two exceptions. These results imply that the patterns of estimation error made by the elementary school and undergraduate groups seem alike when they estimated the measurements of small-sized objects, although the performance of the undergraduates was superior to that of the elementary school groups.

There was a greater frequency of overestimations than underestimations among the fifth graders (volume) and sixth graders (length) for the small-sized TBEOs. Overestimations of length measurements by the sixth graders may possibly result from difficulty with the processing of unit iterations. Processing of unit iterations may lead to estimation errors because of inconsistent unit size or overlapping of units (Joram et al., 2005). Overestimations of volume by the fifth graders may have resulted from the use of multiplication

computation errors. Forrester et al. (1990) suggested that using multiplication computations for estimating area and volume measurements may lead to overestimates. However, further investigations into strategies used by elementary school students in measurement estimation are necessary to clarify the reasons behind such estimation errors.

### Good Estimators' Strategies Used for Estimating Measurements

Good estimators across grade groups described a preference for the use of body parts and nearby objects for estimating the various measurements of school objects in the study. The instruction given prior to task completion recommended the use of body parts or objects in convenience for more accurate measurement estimations (Ministry of Education, 2010), and as such, may have biased the participants.

A few good estimators in the elementary group used the strategy of guessing, which has previously been noted as a less effective strategy (Montague & van Garderen, 2003). As revealed in the interviews, some elementary school students considered guessing when they encountered the following cases: (1) an estimated answer that was not precise, and (2) an estimated outcome obtained by using an inexact length as a measure unit. Additionally, these younger students used other ineffective strategies (e.g., unclear and blank). In contrast, most of the good estimators in the undergraduate group tended to use more effective strategies, such as referencing of body parts, mental rulers, and objects, instead of guessing. The use of more effective strategies may have led to better performance among undergraduates than among the younger groups. The findings imply that enhancements of knowledge of linear size and scale through measurement activities, which is suggested for improving the efficiency of estimation strategy (Tretter et al., 2016), need to be taken into account for strengthening elementary school students' ability to estimate measurements.

Previously, Huang (2016) found that approximately 12% of fifth and sixth graders ( $N = 948$ , 483 fifth graders and 465 sixth graders) used ineffective strategies, while only 5% (20/360) of the fifth- and sixth-grade participants who were identified as good estimators in the current study did so. These two samples differed in terms of estimator characteristics; that is, the current sample was identified as good estimators while the previous sample represented general estimators but not good estimators. Such a distinction in estimation characteristics between the two samples may have led to the differences in the percentage of using ineffective strategies. This implies that the use of effective strategies is associated with estimators' performance of measurement estimation.

Approximately half of the interviewees in the current study described using mental measurement units (e.g., 1 cm, 10 cm, 1 m,  $1 \text{ cm}^2$ , and  $1 \text{ m}^2$ ) that they had constructed from previous experience of measurement for making comparisons when they used eyeballing. Mentally constructing measurement units and making good use of them as references facilitate estimation accuracy (Tretter, Jones, & Minogue, 2006). These results suggest that one essential skill in good estimation performance is the use of eyeballing in combination with mental rulers.

During the interviews, the interviewees mentioned the use of mental rulers, previous experience, and measurement formulas for estimations, but these strategies were not clearly described in their written assessments. For example, results from the written portion of the undergraduate assessment did not indicate the use of measurement formulas or previous experience. The participants may have simplified their descriptions when answering the written assessment; thus, future studies on strategies used for appropriate estimations should include interviews with participants.

Collectively, the current findings suggest that an increase in grade is associated with an increase in measurement estimation ability. Generally, students performed best in estimating the measurements of medium-sized TBEOs, while estimations of small TBEOs were more accurate than those of large TBEOs. A tendency to underestimate the measurements of the large-sized TBEOs was observed in all groups. Surprisingly, the sixth graders tended to overestimate the length of small-sized TBEOs, while the fifth graders were inclined to overestimate the volume of small-sized TBEOs. Those with good estimation abilities showed a preference for using body parts and convenient objects as references for their estimations. In addition, the integration of measurement units constructed from previous experiences with eyeballing was an essential skill used by good estimators. Findings from the current study should be considered when developing measurement curricula, specifically at the elementary school level.

### Limitations and Future Directions

This study had several limitations that impact its generalizability. The majority of TBEOs in the current study were touchable. Reasonable estimates can be obtained by using body parts as measurement tools with touchable objects. Future studies using untouchable TBEOs should be conducted to examine performance in mentally determining the measurements of objects and patterns of estimation error when making measurement estimations. Due to the lack of assessment of the volume estimation measurement of large-sized objects, whether there is a tendency to underestimate the volume of large-sized objects, as was found for length and area, is unknown.

Despite the limitations of the current study, the results suggest that experiences with physical measurements facilitate construction of measurement units as references and aid skills in estimating measurements of objects. Investigating how students who are skilled in measurement estimations select effective strategies over less effective strategies (e.g., guessing) for estimating measurements of objects is a promising area for future work.

## ACKNOWLEDGEMENTS

This research was supported by a grant from the Ministry of Science and Technology (MOST 102-2511-S-845-003-MY3). The author would like to thank reviewers for their helpful suggestions and the teachers of the participating elementary schools as well as the assistances for their supports for the research.

## REFERENCES

- Bright, G. W. (1979). Estimating physical measurement. *School Science and Mathematics*, 79(8), 581-586. <https://doi.org/10.1111/j.1949-8594.1979.tb13900.x>
- Coburn, T. G. (1987). Estimation and mental computation: Measurement estimation-making adjustments. *Arithmetic Teacher*, 34(7), 24-25. <https://doi.org/10.5951/AT.34.7.0024>
- Desli, D., & Giakoumi, M. (2017). Children's length estimation performance and strategies in standard and nonstandard units of measurement. *International Journal of Research in Mathematics Education*, 7(3), 61-84.
- Forrester, M. A., & Shire, B. (1994). The influence of object size, dimension and prior context on children's estimation ability. *Educational Psychology*, 14(4), 451-465. <https://doi.org/10.1080/0144341940140407>
- Forrester, M. A., Latham, J., & Shire, B. (1990). Exploring estimation in young primary school children. *Educational Psychology*, 10, 283-300. <https://doi.org/10.1080/0144341900100401>
- Gooya, Z., Khosroshahi, L. G., & Teppo, A. R. (2011). Iranian students' measurement estimation performance involving linear and area attributes of real-world objects. *ZDM The International Journal on Mathematics Education*, 43, 709-722. <https://doi.org/10.1007/s11858-011-0338-1>
- Hildreth, D. J. (1983). The use of strategies in estimating measurements. *Arithmetic Teacher*, 30, 50-54. <https://doi.org/10.5951/AT.30.5.0050>
- Huang, H.-M. E. (2016). Investigation of the performance and use of strategy among elementary school children in estimating measurement of length, area, and volume. *Journal of Research in Education Science*, 61(3), 131-162. (In Chinese)
- Huang, H.-M. E., Heinze, A., Ruwisch, S., Hoth, J., & Chang, H.-W. (2019). Investigating grade 7-9 students' performance of length estimation. In A., Rogerson, & J. Morska (Eds.) (2019). *Proceedings of the 15th International Conference of Theory and practice: An interface or a great divide?* (pp. 223-228). Kildare, Ireland: Maynooth University.
- Jones, G., & Taylor, A. (2009). Developing a sense of scale: Looking backward. *Journal of Research in Science Teaching*, 46(4), 460-475. <https://doi.org/10.1002/tea.20288>
- Jones, G., Taylor, A., & Broadwell, B. (2009). Estimating linear size and scale: Body rulers. *International Journal of Science Education*, 31(11), 1495-1509. <https://doi.org/10.1080/09500690802101976>
- Jones, M. G., Forrester, J. H., Gardner, G. E., Andre, T., & Taylor, A. R. (2012). Students' accuracy of measurement estimation: Context, units, and logical thinking. *School Science and Mathematics*, 112(3), 171-178. <https://doi.org/10.1111/j.1949-8594.2011.00130.x>
- Jones, M. G., Tretter, T., Taylor, A., & Oppewal, T. (2008). Experienced and novices teachers' concepts of spatial scale. *International Journal of Science Education*, 30(3), 409-429. <https://doi.org/10.1080/09500690701416624>
- Joram, E., Gabriele, A., Bertheau, M., Gelman, R., & Subrahmanyam, K. (2005). Children's use of the reference point strategy for measurement estimation. *Journal for Research in Mathematics Education*, 36(1), 4-23.
- Joram, E., Subrahmanyam, K., & Gelman, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research*, 68, 413-419. <https://doi.org/10.3102/00346543068004413>
- Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 179-192). Reston, VA: National Council of Teachers of Mathematics.
- Ministry of Education (2010). *Grade 1-9 curriculum for junior-high school and elementary school: Mathematics* (3rd ed.). (In Chinese.) Taipei, Taiwan: The Author.
- Ministry of Education (2018). *Curriculum guidelines of 12-year basic education: Mathematics*. (In Chinese.) Taipei, Taiwan: The Author.
- Montague, M., & Van Garderen, D. (2003). A cross-sectional study of mathematics achievement, estimation skills, and academic self-perception in students of varying ability. *Journal of Learning*

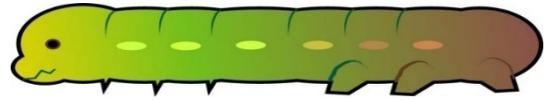
- Disabilities*, 36(5), 437-488. <https://doi.org/10.1177/00222194030360050501>
- National Center for Educational Statistics (2013). NAEP questions (Question ID 2013-4M3 #10 M1461E1). Retrieved from National Center for Educational Statistics Website: <https://nces.ed.gov/NationsReportCard/nqt/Search>
- National Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Reston, VA: The Author.
- Pike, C. D., & Forrester, M. A. (1997). The influence of number-sense on children's ability to estimate measures. *Educational Psychology*, 17(4), 483-500. <https://doi.org/10.1080/0144341970170408>
- Ruwisch, S., & Huang, H.-M. E. (2018). Length measurement and estimation in primary school- A comparison of the curricula of Taiwan and Germany. In F.-J. Hsieh (Ed.). *Proceedings of the 8<sup>th</sup> ICMI-East Asia Regional Conference on Mathematics Education* (Vol. 2, pp. 269-280). Taipei, Taiwan: National Taiwan Normal University.
- Ruwisch, S., Heid, M., & Weiher, D. F. (2015). Children's uses of strategies in estimating length and capacity. In Beswick, K., Muir, T., & Wells, J. (Eds.). *Proceedings of 39th Psychology of Mathematics Education Conference* (Vol. 1, pp. 246). Hobart, Australia: PME.
- Ryan, J., & Williams, J. (2007). *Children's mathematics 4-15: Learning from errors and misconceptions*. Berkshire, England: Open University Press.
- Sternberg, R. J., & Sternberg, K. (2017). *Cognitive psychology* (7th ed.). Boston, MA: Cengage Learning.
- Swan, M., & Jones, O. (1980). Comparison of students' percepts of distances, weight, height, area, and temperature. *Science Education*, 64, 297-307. <https://doi.org/10.1002/sce.3730640305>
- Tretter, T. R., Jones, M. G., & Minogue, J. (2006). Accuracy of scale conceptions in science: Mental maneuverings across many orders of spatial magnitude. *Journal of Research in Science Teaching*, 43(10), 1061-1085. <https://doi.org/10.1002/tea.20155>
- Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching*, 43(3), 282-319. <https://doi.org/10.1002/tea.20123>

## APPENDIX A

### Measurement Estimation Assessment

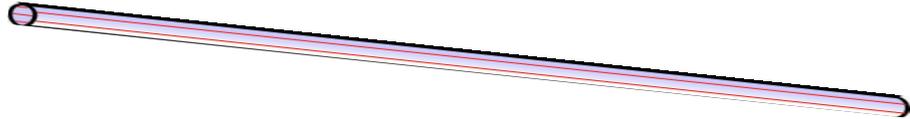
1. Approximately how long is the caterpillar shown below? (Do not use a ruler).

( ) cm



- ( ) 2. Approximately how many  $1 \text{ cm}^3$  cubes can be lined up along the straw shown below without any remaining space or overlap?

(Ⓐ 9; Ⓑ 12; Ⓒ 15; Ⓓ 18).



- ( ) 3. About how many  $1 \text{ cm}^3$  cubes would it take to construct the small eraser without any remaining space or overlap?

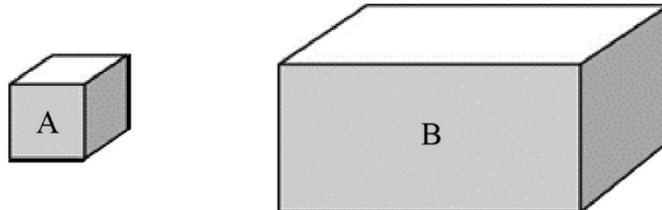
(Ⓐ More than 6 but fewer than 7; Ⓑ More than 8 but fewer than 9; Ⓒ More than 10 but fewer than 11).

- 4-1. What is the approximate volume of the large colored chalk box?

( )  $\text{cm}^3$

- 4-2. My estimation method(s) :

5. Figure A is a cube with a side of 1 cm.



- 5-1. What is the approximate volume of the box B?

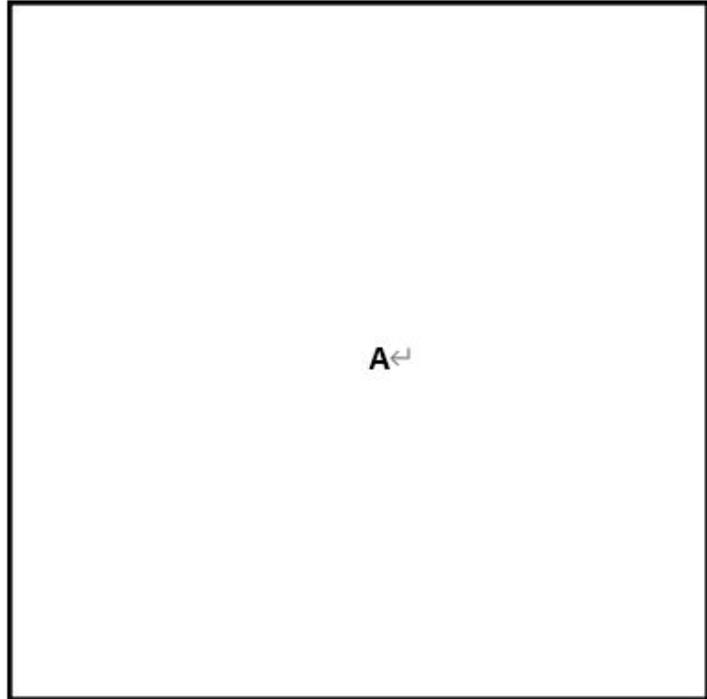
( )  $\text{cm}^3$

- 5-2. About how many cubes of 2 cm in length, 1 cm in width and 1 cm in height, would it take to construct box B without any remaining space or overlap? ( )

- ( ) 6. What is the approximate volume of the white-chalk box?

(Ⓐ More than  $800 \text{ cm}^3$  but less than  $900 \text{ cm}^3$ ; Ⓑ More than  $1,000 \text{ cm}^3$  but less than  $1,100 \text{ cm}^3$ ; Ⓒ More than  $1,200 \text{ cm}^3$  but less than  $1300 \text{ cm}^3$ ).

7. What is the approximate area of figure A?  
 ( )  $\text{cm}^2$



- 8-1. What is the approximate area of this piece of paper?  
 ( )  $\text{cm}^2$

8-2. My estimation method(s) :

9. Which object in the classroom has an area closest to  $1 \text{ m}^2$ ? Please write it down.

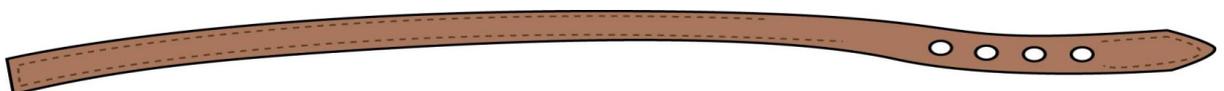
- ( ) 10. What is the approximate surface area of the garbage bag?  
 (① More than  $1,400 \text{ cm}^2$  but less than  $1,700 \text{ cm}^2$ ; ② More than  $1,800 \text{ cm}^2$  but less than  $2,100 \text{ cm}^2$ ; ③ More than  $2,200 \text{ cm}^2$  but less than  $2,500 \text{ cm}^2$ ).
- ( ) 11. About how many pieces of 8K drawing paper would it take to fully cover the piece of plastic sheet with an area of  $5,260 \text{ cm}^2$  without any remaining space or overlap?  
 (① More than 5 but fewer than 5.5; ② More than 6 but fewer than 6.5; ③ More than 7 but fewer than 7.5).

- 12-1. Approximately how long is the string? (Do not use a ruler).  
 ( ) cm

12-2. My estimation method(s) :

13. Please draw a straight line of about 25 cm. (Do not use a ruler).

- ( ) 14. About how many pieces of chalk are 54 cm long? (① 5; ② 7; ③ 9; ④ 11).
- ( ) 15. About how many pieces of 3-cm ribbon would it take to lay along the belt shown below without any remaining space or overlap?  
 (① More than 3 but fewer than 3.5; ② More than 4.5 but fewer than 5; ③ More than 6 but fewer than 6.5).



16. Please draw a rectangle (or square) with an area of  $16 \text{ cm}^2$  (Do not use a ruler).

## APPENDIX B

## Characteristics of the Measurement Estimation Assessment: Size Section, Measurement Attribute, Item Format, Types of TBEO Presentation, Size of TBEOs and Measure Unit (or Benchmark Prescribed), and Answer

Size section	Item number and question	Measurement attribute	Item format	Presentation	To-be-estimated object [TBEO], measure unit (or benchmark prescribed) and answer
Small	Q1. Estimate the length of the caterpillar	Length	Fill-in	Picture	1. TBEO (the length of the caterpillar) = 7 cm Answer : 7 cm
	Q7. Estimate the area of figure A	Area	Fill-in	Picture	1. TBEO (the area of figure A) = $[10 \times 10 = 100 \text{ (cm}^2\text{)}]$ Answer : 100 cm <sup>2</sup>
	Q16. Draw a rectangle (or a square) with an area of 16 cm <sup>2</sup>	Area	Drawing	Others	A drawn rectangle or square with an area 6 cm <sup>2</sup>
	Q3. About how many 1 cm <sup>3</sup> cubes would it take to construct the eraser	Volume	Multiple-choice	Real object	1. TBEO (the volume of the eraser) = $4.4 \times 1.2 \times 1.7 = 8.98 \text{ (cm}^3\text{)}$ ; 2. Measure unit = 1 cm <sup>3</sup> Answer : ② More than 8 but fewer than 9
	Q5-1. Estimate the volume of box B	Volume	Fill-in	Real object	1. TBEO (the volume of box B) = $[4 \times 2 \times 2 = 16 \text{ (cm}^3\text{)}]$ Answer : 16 cm <sup>3</sup>
	Q5-2. About how many blocks would it take to construct box B	Volume	Fill-in	Real object	1. TBEO (the volume of box B) = $[4 \times 2 \times 2 = 16 \text{ (cm}^3\text{)}]$ ; 2. Benchmark prescribed (blocks) = $[2 \times 1 \times 1 = 2 \text{ (cm}^3\text{)}]$ Answer : 8 blocks
Medium	Q2. About how many 1 cm <sup>3</sup> blocks can be lined up as long as the straw	Length	Multiple-choice	Picture	1. TBEO (the length of the straw) = 12 cm 2. Measure unit : 1 cm <sup>3</sup> Answer : ②12
	Q13. Draw a line of about 25 cm	Length	Drawing	Others	A Drawn a line of about 25 cm
	Q15. Estimate the length of the belt	Length	Multiple-choice	Picture	1. TBEO (the length of the belt) = 19 cm 2. Benchmark prescribed (ribbon) : 3 cm Answer : ③More than 6 but fewer than 6.5
	Q4-1. Estimate the volume of the big colored chalk box	Volume	Fill-in	Real object	1. TBEO (the volume of the large colored chalk box) = $[5 \times 8.5 \times 12.8 = 544 \text{ (cm}^3\text{)}]$ Answer : 544 cm <sup>3</sup>
	Q6. Estimate the volume of the white-chalk box	Volume	Multiple-choice	Real object	1. TBEO (the volume of the white-chalk box) = $[12 \times 9 \times 8 = 864 \text{ (cm}^3\text{)}]$ Answer : ①More than 800 cm <sup>3</sup> but less than 900 cm <sup>3</sup>
Large	Q12-1. Estimate the length of the string	Length	Fill-in	Real object	1. TBEO (the length of the string) = 78 cm Answer : 78 cm
	Q14. About how many chalks is 54 cm	Length	Multiple-choice	Real object	1. TBEO (the length of the string) = 54 cm 2. Benchmark prescribed (chalks) = 8 cm Answer : ② 7 chalks
	Q8-1. Estimate the area of the piece of paper	Area	Fill-in	Real object	1. TBEO (the area of the piece of paper) = $[30 \times 50 = 1500 \text{ (cm}^2\text{)}]$ Answer : 1500 cm <sup>2</sup>
	Q9. Which object in the classroom has an area closest to 1 m <sup>2</sup>	Area	Fill-in	Others	List one classroom object with an area closest to 1 m <sup>2</sup> of objects
	Q10. Estimate the area of the garbage bag	Area	Multiple-choice	Real object	1. TBEO (the area of the garbage bag) = $[43 \times 56 = 2408 \text{ (cm}^2\text{)}]$ Answer : ③More than 2200 cm <sup>2</sup> but less than 2500 cm <sup>2</sup>
	Q11. About how many pieces of 8K drawing paper would it take to fully cover the area of 5260 cm <sup>2</sup>	Area	Multiple-choice	Real object	1. TBEO (the area of the plastic-sheet) = 5260 cm <sup>2</sup> 2. Benchmark prescribed = 8K drawing paper $[39.5 \times 27 = 1066.5 \text{ (cm}^2\text{)}]$ Answer : ① more than 5 but fewer than 5.5

Note. 1. Fill-in = Fill-in-the-blank. 2. Others = Neither a TBEO nor a benchmark prescribed was presented.