ISSN 2234-3075(Print)
ISSN 2288-7806(Online)
Http://dx.doi.org/10.14771/AIM.8.2.4

Research on the User-specific Visualization of Data with Graphs in the Agricultural Sector

Sohyun In[1], Chunghan Kang[2], Jihye You[3], Dohyeong Choi[4], Junghoon Moon[5]

This research is about visualizing data for efficient use in agriculture, by measuring the understanding of graphs according to their features. A questionnaire consisting of one graph, either a bar graph or line graph, and a question about data, was made. According to the results, the time spent answering the question did not differ between the line graph and the bar graph. However, the rate of correct answers differed: that of the bar graph was 50% and the line graph was 0%. This implies that participants answered using the bar graph more accurately regarding whether the data involved understanding the absolute value. From this result, it is clear that using proper features of graphs according to the characteristics of data improves the understanding of information. In addition, a strategy to utilize information and communication technology in agriculture effectively was discussed.

1 First author, Food biz lab, Master Student,
Dept of Agricultural Economics & Rural Development,
Seoul National University
sohyun6241@snu.ac.kr

2 Food biz lab, Ph.D Student,
Dept of Agricultural Economics & Rural Development,
Seoul National University
ggulbbang@gmail.com

3 Food biz lab, Master Student,
Dept of Agricultural Economics & Rural Development,
Seoul National University
jhyou2399@snu.ac.kr

4 Food biz lab, Undergraduate Student,
Dept of Agricultural Economics & Rural Development,
Seoul National University
ehgudchl@snu.ac.kr

5 Food biz lab, Associate Professor Dept of Agricultural Economics & Rural Development, Seoul National University
moonj@snu.ac.kr(⌧)

Article history
Received 7 November 2016
Received in revised form 9 November 2016
Accepted 11 November 2016

1. Introduction

Agriculture in Korea has been struggling with diminished labor forces due to aging in rural society (Lee, Hwang, and Yoe, 2014) and faced with significant tasks, as it has to strengthen its international competitive power in the world market (Moon and Hwang, 2014). To overcome these challenges, information and communication technologies (ICT) in agriculture has been newly recognized as a high value-added business since Korean rural society can utilize internal advanced IT and its infrastructure to substitute old labor forces and improve their productivity. As the introduction of these technologies has been supported strongly by related government departments, increasing numbers of farms have been applying ICT and related research has been actively conducted (Kim et al., 2013).

In Korea, ICT has been mainly introduced in horticulture. Many farms and rural societies, however, have difficulties in creating new values and improving their productivity by introducing ICT. According to Shim et al. (2014), the improper use of measured data is the most significant problem with making full use of ICT in horticulture.

The use of real-time databases is a major factor in controlled horticulture. There are environmental and control data, such as internal air temperature and internal CO2, and growth data, such as the leaf length and stem diameter, in these databases. These stored data are used to construct the optimal growth models for each crop; therefore, the use of models has great effects on growth, output, and quality of crops (Hur et al., 2011). Moreover, by using the data, rural consultants can provide data-based consulting services to their customers instead of unscientific experience-based services, and all these performances create added value in agriculture (Shim et al., 2014).

Inadequate data use in agriculture is deeply related to the feature of agricultural management. The data must be considered simultaneously in a prudent manner to make farm management decisions since the yield, quality, and growth of crops are determined by diverse and complex environmental conditions (Hur et al., 2011). Thus, multivariate graphs should be used in the practical stages, such as farm diagnosis or management consultation. However, they limit intuitive recognition of the graphs and create biases that making people confused to recognize information.

In this regard, the purpose of this study is to determine a suitable graphical form that enables farm managers to use measured data appropriately without bias. We anticipate that this research will provide us with reasonable and systemic guidelines of ICT in horticulture to help farm managers recognize the environmental, controlling, and crops' growth data from a greenhouse efficiently and effectively.

According to previous research (Galesic and Garcia-Retamer 2011, Yasima et al., 2011) insisting that the ways in which people acquire information differ based on the features of a graph, it proved whether the recognition of graphs differed using different features of the graph in same information. This research studied the relationship between the understanding of graphs and features of graphs. Time and accuracy of recognizing information from graphs were used to measure the effectiveness and efficiency of the understanding of graphs. Different features of graphs were given to users including the same information. From these graphs, their accuracy and time taken for recognition were measured.

2. Theoretical Background

2.1 Graph literacy

Graph literacy is "the ability to understand graphically presented information" (Galesic and Garcia-Retamero, 444, 2011). Since the use of visual displays of quantitative data is pervasive in our highly technological society (Friel, 2001) and in the last 10 years, the usage of graphs in academic journals and newspapers has more than doubled (Zacks et al., 2002), graph literacy has become important for people to recognize and use given information. As graphs can include more information than words and send messages efficiently, visual displays such as line plots or bar charts facilitate the communication of information by enabling the representation of quantitative information in spatial locations (Okan et al., 2012). Graph literacy appears differently in every person, so people with high graph literacy are able to recognize the elaborate relationship of information through a graph, such as understanding trends in information through a bar graph (Okan et al., 2012). This ability differs among individuals by numeracy, age, or people's familiarity with the specific content depicted (Okan et al., 2012). Since graphs are widely used, it is important to compose so that even people with relatively low graph literacy can understand information through the graph.

2.2 Understanding of a graph according to its features

A graph is one way of transmitting information based on the position of a point, line, or area on a two-dimensional surface (Fry, 1983). It consists of 4 components: the framework, specifier, label, and background (Friel, 2001). In addition, there are many kinds used, such as bar graphs, histograms, line graphs, scatterplots, and pie charts. As the features of graphs vary, some graphs are better at conveying certain kinds of information than others (Lohse, 2009). It is known that viewers are faster at reading individual data points when viewing bar graphs compared to line graphs, and they are faster at making trend judgments when viewing line graphs compared to bar graphs (Simcox, 1984). In addition, viewers can more accurately identify individual data points from bar graphs than from line graphs (Shah and Freedman, 2011). When presenting information in bar graphs, categorical information is usually used. That is, discrete data are proper for bar graphs, as they are easy to use for comparing categories, such "higher," "lower," "greater than," and "less than." On the other hand, information presented as line graphs is mainly described as continuous trends between the data points, using terms such as "rising," "falling," "increasing,'" and "decreasing" (Zacks and Tversky, 1999). Because of these characteristics, people usually recognize that the information in bar graphs is discrete, and information in line graphs is continuous. This implies that the features of graphs are important in transferring information to individuals. Previous studies have implied that their features significantly impact humans' recognition of information through graphs.

3. Methodology

3.1 Research model and hypothesis

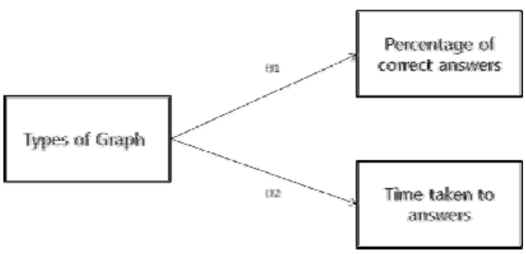
People can easily understand a concept when they read information shown as a graph (Macdonald-Ross, 1977). A graph can be a useful tool when it illustrates a quantitative concept clearly in the display (Larkin, 1987). However, showing graphs without considering the characteristics of the data leads to errors for readers (Shah and Carpenter, 1995). Therefore, it is important to use proper graphs depending on the types of data. According to Galesic and Garcia (2011), people can read the absolute value of data shown in a bar graph. In contrast, people can read the stream of data better when displayed by a line graph.

Time is one of the dimensions assessing how participants understand data (Shah and Hoeffner, 2002). And one of methods to verify performance of graphs is taking the task such as finding important element on graphs (McCaffery et al., 2012). Based on the above information, this study established the following hypotheses:

H1: The percentage of correct answers from discrete data should be higher in a bar graph than a line graph.
H2: Less time should be required to provide a correct answer for discrete data shown in a bar graph than a line graph.

This study aims to clarify the differences in the percentage of correct answers and answering time depending on the types of graph. The research model of this study is shown in Figure 1.

Figure 1 Research model



3.2 Method

To verify the hypotheses, a total of 16 horticulture farm owners participated. Subjects were given either a bar graph or line graph randomly. Both graphs have data and contents. All participants heard an explanation of graphs. Both graphs show the stem diameter by internal air temperature and CO_2. Figure 2 shows a bar graph, and Figure 3 shows line graph participants had viewed. Temperature and CO_2 are both important discrete elements of a plant's stem growth (Peltola, Kilpeläinen, and Kellomäki, 2002). Then, participants were asked to answer an accompanying question about which one is a more important element of the stem diameter's growth. Table 2 shows the question participants were given. The percentage of correct answers and the time taken to answer were analyzed.

Table 1 shows the demographic information. All the subjects in this study were men. Their average age was 48.1. They raised crops such as strawberries, tomatoes, paprika and green pumpkin for about 12 years.

Table 1 Demographic information

Category		Mean	Count	%
Sample			16	100
Sex	Male		16	100
	Female		0	0
Age		48.1		
Location	Jeolla-do		12	75
	Gyeongsang-do		3	18.8
	Busan		1	6.2
Farming experience		12		
Type	Strawberry		6	37.5
	Tomato		5	31.3
	Paprika		4	25
	Green pumpkin		1	6.2

Figure 2 Bar graph of the stem diameter according to internal air temperature and CO_2

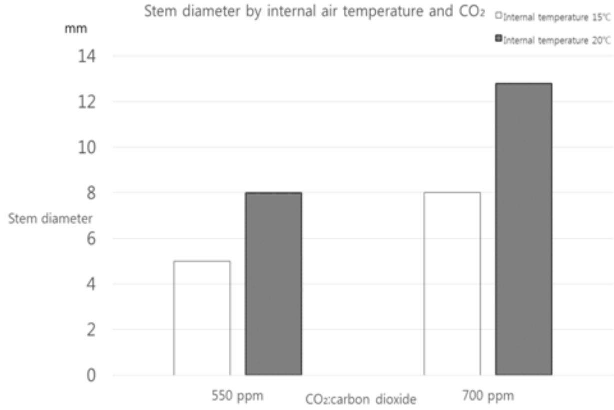


Figure 3 Line graph of the stem diameter according to internal air temperature and CO_2

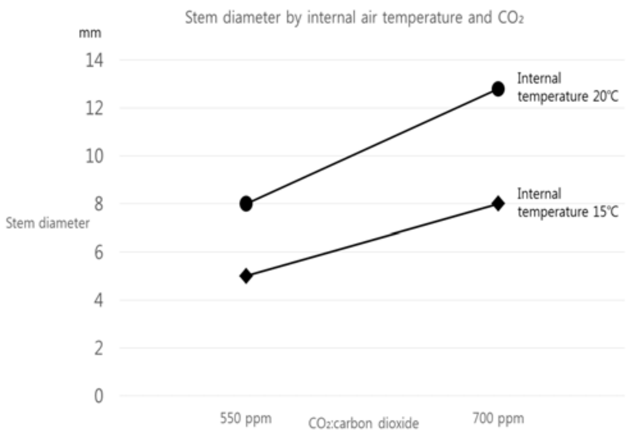


Table 2 Question comparing the bar and line graphs

Q: On the graph, which is a more important element of the stem diameter's growth?
1. CO_2 2. Internal temperature 3. Both 4. Don't know

4. Statistical Analysis and Results

The data on the 16 farmers included in the analysis were aggregated and analyzed using a t-test. The results are listed in Table 3. 7 subjects were tested using the bar graph, and 9 were tested using the line graph. The results show that the average duration of testing was not affected by the different graph types ($t=-1.3997$, $p>0.05$), but in terms of the percentage of correct answers, there was a difference based on the graph type. 3 of 7 participants who were tested using the bar graph obtained correct answers, whereas no one tested using the line graph answered the same question correctly. Consequently, the type of graph had no effect on the answering time duration, but it did affect the percentage of correct answers.

Table 3 The t-test results

	Average testing time		
	Bar graph	Line graph	t-value
Means	63.71	46.50	-1.3997

5. Discussion

This study has investigated the impact of graph types on the answering duration and percentage of correct answers. The results show that participants recognized data more accurately when illustrated in a bar graph. It might be useful to compare previous studies, like those of Padilla, McKenzie and Shaw (1986) Friel, Curcio and Bright (2001), Shah, Mayer, and Hegarty (1999), which investigated graph literacy in students, as this study included farmers who are relatively older than the average ages of students. Elderly people tend to have lower reading and numeracy skills and comprehension (Oldfield and Dreher, 2010). In addition, elderly people have had few opportunities to be exposed to a variety of types of graphs. As this research showed different recognition for farmers based on different features of graphs, the findings can serve as a guideline for the effective use of graphs for farmers.

This research suggests the using proper graphs to print out data by characteristics of information in the agriculture sector. As this research showed the different literacy in different kinds of graph, therefore, this research gives the guideline of visualizing data in agriculture which starts using ICT technology. For instance, when continuous data, such as time series data and daily environmental data, are given, they should be presented in a line graph. Moreover, data on discrete categories that need to show absolute values should be shown in a bar graph.

This study has several limitations. It included 16 participants, which is a relatively small number; therefore, the results were not clear. If the number of participants was increased, the results would have been clearer. In addition, the experiment used only discrete data for the test, which are known to be appropriate for a bar graph. The finding in which the bar graph had a higher rate of correct answers stemmed from this condition. Use of discrete data and continuous data equally would be better to support the thesis that insists importance of matching proper kinds of graph and data.

Further research should ask participants 2 questions regarding discrete data and continuous data. Since previous studies have insisted that the characteristics of data determine which type of graph to use, it would be preferable to make questions set to acquire different types of data. In addition, more samples are needed in further research to generalize the results and make more confident guidelines.

Since farmers had not had frequent exposure to the graphs, guideline for describing growth data, environment data on greenhouses, and control information on greenhouses are needed. The present findings showed that proper features of graphs are needed to ensure that farmers understand data effectively. In addition, this result is the starting point for creating guidelines in visualizing information in agriculture, focusing on farmers.

Acknowledgement
This research was supported by the Ministry of Science, ICT and Future Planning (MSIP), Korea, under the Information Technology Research Center (ITRC) support program (IITP-2016-H8601-16-1007) supervised by the Institute for Information & Communications Technology Promotion (IITP).

References

M. Galesic and R. Garcia-Retamero (2011). Graph literacy a cross-cultural comparison. *Medical Decision Making*, 31(3), pp.444-457.

H.A.S. Jill H. Larkin (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11(1), pp.65-100.

M. Macdonald-Ross (1977). Graphics in Texts. *Review of Research in Education*, 5(pp.49-85).

K.J. McCaffery, A. Dixon, A. Hayen, J. Jansen, S. Smith and J.M. Simpson (2012). The Influence of graphic display format on the interpretations of quantitative risk information among adults with lower education and literacy a randomized experimental study. *Medical Decision Making*, 32(4), pp.532-544.

P. Shah and P.A. Carpenter (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General,* 124(1), pp.43.

P. Shah and J. Hoeffner (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review,* 14(1), pp.47-69.

Zacks, J., Levy, E., Tversky, B., & Schiano, D. (2002). Graphs in print. In M. Anderson, B. Meyer, & P. Olivier (Eds.), Diagrammatic representation and reasoning (pp. 187–206). London: Springer-Verlag.

S.N. Friel, F.R. Curcio and G.W. Bright (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education,* pp.124-158.

M. Galesic and R. Garcia-Retamero (2011). Graph literacy: a cross-cultural comparison. *Med Decis Making*, 31(3), pp.444-57.

H.A.S. Jill H. Larkin (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11(1), pp.65-100.

G.L. Lohse (2009). A Cognitive Model for Understanding Graphical Perception. *Human–Computer Interaction*, 8(4), pp.353-388.

M. Macdonald-Ross (1977). Graphics in Texts. *Review of Research in Education*, 5(pp.49-85.

K.J. McCaffery, A. Dixon, A. Hayen, J. Jansen, S. Smith and J.M. Simpson (2012). The Influence of graphic display format on the interpretations of quantitative risk information among adults with lower education and literacy a randomized experimental study. *Medical Decision Making*, 32(4), pp.532-544.

Y. Okan, R. Garcia-Retamero, E.T. Cokely and A. Maldonado (2012). Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension. *Journal of Behavioral Decision Making*, 25(4), pp.390-401.

S.R. Oldfield and H.M. Dreher (2010). The concept of health literacy within the older adult population. *Holistic nursing practice*, 24(4), pp.204-212.

M.J. Padilla, D.L. McKenzie and E.L. Shaw (1986). An examination of the line graphing ability of students in grades seven through twelve. *School Science and Mathematics,* 86(1), pp.20-26.

H. Peltola, A. Kilpeläinen and S. Kellomäki (2002). Diameter growth of Scots pine (Pinus sylvestris) trees grown at elevated temperature and carbon dioxide concentration under boreal conditions. *Tree Physiology*, 22(14), pp.963-972.

P. Shah and P.A. Carpenter (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General*, 124(1), pp.43.

P. Shah and E.G. Freedman (2011). Bar and line graph comprehension: an interaction of top-down and bottom-up processes. *Top Cogn Sci,* 3(3), pp.560-78.

P. Shah and J. Hoeffner (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review,* 14(1), pp.47-69.

P. Shah, R.E. Mayer and M. Hegarty (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology,* 91(4), pp.690.

W.A. Simcox (1984). A method for pragmatic communication in graphic displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 26(4), pp.483-487.

F.R.C.a.G.W.B. Susan N. Friel (2001). Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education,* 32(2), pp.124-158.

J. Zacks and B. Tversky (2013). Bars and lines: A study of graphic communication. *Memory & Cognition,* 27(6), pp.1073-1079.

J.W Lee, J.H Hwang, and H. Yoe (2014). The trend and direction of ICT Convergence Technology in Korean agriculture. *Information & Communications Magazine,* 31(5), pp.54-60

J.H Moon, and Y. M Hwang (2014). The latest cases of ICT in Korean agriculture and food business. I*nformation & Communications Magazine,* 31(5), pp.61-66'

Y.J Kim, S.Y Gouk, Y.R Kim, M.G Kim, J.S Kim, and Y.H Kim (2013). The Present status and Development Direction of Smart Agriculture. *Korean Rural Economic Institute Policy Study Report*, pp.1-159

K.S Shim, H.R Lee, M. Kim and J.H An (2014). The Realization Plan of a Controlled Horticulture Smart Farm based ICT Convergence. *Journal of the Korea society of IT services: collection of dissertations on academic conference*, pp. 485-488.

S.O Hur, K.H Han, S.H Jeon, Y.S Jang, S.W Kang, S.O Chung, H.J Kim and K.H Lee (2011) Application of smartphone and wi-fi communication for remote monitoring and control of protected crop production environment. *CNU Journal of Agricultural Science* 38(4), pp. 753-759