

Physico-Symbolic Characteristics of the Japanese Paperback Book Series *Shinsho*: A Descriptive Study

Shohei Yamada

The University of Tokyo,
Graduate School of Education
syamada@p.u-tokyo.ac.jp

Ruri Shimura

The University of Tokyo,
Graduate School of Education
shimshim_rr@hotmail.co.jp

Bin Umino

Toyo University,
Faculty of Sociology
umino@toyo.jp

Shin'ichi Toda

Toyo University,
Faculty of Sociology
toda@toyo.jp

Kyo Kageura

The University of Tokyo,
Interfaculty Initiative in Information Studies
kyo@p.u-tokyo.ac.jp

ABSTRACT

Background. Printed books occupy a privileged position in the knowledge infrastructure of modern society. Nevertheless, their physical and symbolic arrangements and characteristics have not yet been systematically studied.

Objectives. We observed several physico-symbolic characteristics of the Japanese paperback book series *Shinsho* and revealed how they changed over the past 70 years. We selected *Iwanami Shinsho* for analysis, as it has the longest history among the current *Shinsho* series and is widely held to be the best *Shinsho* series in providing lay readers with authentic knowledge.

Methods. We analysed the physico-symbolic characteristics of the books, including size and thickness of the books, size of live area, and arrangement of characters, and then provided descriptive statistics and plots for these characteristics.

Results. This study revealed some interesting points related to chronological change. First, the height of the live area has increased by 7 mm in the past 70 years, even though the height of the pages has not changed so much. Second, the

number of characters per double page spread has decreased since the 1970s. Third, while the size of the content was stable at around 140,000 characters from the 1950s to the 1990s, there is a decreasing trend from the 1990s.

Contributions. Understanding how knowledge for lay readers is represented within books is important, given that the form and arrangement of knowledge representation affects how knowledge is understood. In addition, the results of the study provide useful information for the design of the physico-symbolic characteristics of *Shinsho* and other book forms.

INTRODUCTION

Printed books have occupied a unique and privileged position in the representation, distribution, and consumption of knowledge since the invention of movable-type printing (Eisenstein, 1980). Although the importance of printed books is widely recognised and accepted, most of us have not paid much attention to the physical and symbolic characteristics of books and the arrangement of books in society.

The advent of the Internet and especially of e-books, and the recognition that they may pose a threat to the privileged position that printed books have enjoyed so far have revitalised discussion about printed books (Lyons, 2011; Pearson, 2013; Cave & Ayad, 2014). However, most recent discussions about printed books are rather impressionistic and empirical (Merkoski, 2013; Eco & Carrière, 2009/2010). In order to compare printed books and e-books as knowledge representation media, it is essential to review the state of printed books as an established and privileged media at the outset. A detailed characterisation of the physical and symbolic characteristics of books and the arrangement of books in society is one way to represent the actual state of printed books.

Printed books can be classified into socially-consolidated and production-oriented books. Yamada, Shimura, Umino, Toda and Kageura (2015) analysed the physico-symbolic characteristics of Japanese bestsellers, which are a socially-consolidated class of books. We now focus on production-oriented class of books. Production-oriented books take various forms in terms of series, format, and publication. Each form of books has its own role in the representation and distribution of knowledge. Therefore, it is essential to investigate each group of books with similar form in order to comprehend the physical and symbolic characteristics of production-oriented books.

Within this context, we investigated the physico-symbolic characteristics, the functional arrangement, and structure of chapters and sections of the Japanese paperback book series, *Shinsho*. *Shinsho* is one of the popular forms of Japanese paperback books; the basic purpose behind its publication is to provide lay readers with authentic knowledge. The role of *Shinsho* resembles that of Pelican Books (Penguin Books), *Que sais-je?* (Presses Universitaires de France), and the Very Short Introduction series (Oxford University Press).

In Japan, several publishers publish their own *Shinsho* series. The number of *Shinsho* titles published yearly is large.

This paper details the results of a part of a bigger survey, focusing on several physico-symbolic aspects of Japanese *Shinsho*. The following section introduces the target and method of our investigation, together with the characteristics to be investigated. The next section provides summary descriptions, after which we discuss how the physico-symbolic characteristics have changed over time.

DATA AND CHARACTERISTICS

Data: Iwanami Shinsho Series

As explained earlier, there are several publishers that publish their own *Shinsho* series. In recent years, over 1000 titles are published annually as *Shinsho* in Japan (The editorial department of Yearbook of Publication, 2015). Among many *Shinsho* series, we focused on *Iwanami Shinsho* in our study. *Iwanami Shinsho* is the first *Shinsho* series in Japan, starting in 1938. It was published by *Iwanami Shoten Publishers*, one of the biggest academic and scholarly publishers in Japan. *Iwanami Shoten Publishers* has as its main purpose for publishing its *Iwanami Shinsho* series the fostering of public culture; and its authors include famous scholars and experts (The editorial department, *Iwanami Shoten Publishers*, 1988). This makes *Iwanami Shinsho* the best example of a *Shinsho* series. *Iwanami Shinsho* thus constitutes a basic point of departure for analysing *Shinsho* series in general. Imamura (2009) and Imamura (2014) described some physical and symbolic characteristics of *Iwanami Shinsho*, but the analysis remained partial. To the best of our knowledge, there is no study that addresses the physico-symbolic characteristics of paperback books in English.

From its beginning in 1938 to the end of 2014, 3,018 titles were published under the *Iwanami Shinsho* series. This number excludes three revised editions, which have the same serial numbers as the original editions, and ten supplementary volumes. In the case of multi-volume works, we regarded each volume as single title. We sampled from these 3,018 titles. Since one of our aims was to observe the chronological transition of physico-symbolic characteristics, it was necessary that the sample reflected the order of the publication date. Therefore, we sampled the titles for analysis according to the following procedure. First, we listed the 3,018 titles of *Iwanami Shinsho* in order by series number. Since the series number was basically assigned based on the order of the publication date, we selected titles whose series number has 3 as the last digit. Altogether, we selected 302 titles that reflect the order of the publication date. The *Iwanami Shinsho* series has four versions, which are designated by different colours and separated by a certain period of time. These include the red version (from 1938 to 1946), the blue version (from 1949 to 1977), the yellow version (from 1977 to 1987), and the new red version (from 1987 to the present). Table 1 shows the period of publication of each colour version. It also shows the number of titles both for the listed 3,018 titles and the sampled 302 titles.

Physico-symbolic Characteristics

For these titles, we identified, measured and recorded 53 characteristics, classified into five categories:

1. Eight types of bibliographic information: title and subtitle, publication date, authors, editors, translators, price, edition and impression, and ISBN;
2. Ten physical and physico-symbolic aspects of the books: height of content page, width of

Table 1. Period of publication of each colour version

	red	blue	yellow	new red	total
period of publication	1938 - 1946	1949 - 1977	1977 - 1987	1987 - 2014	
number of titles published	101	1,000	396	1,521	3,018
number of titles sampled	10	100	40	152	302

content page, head margin, tail margin, fore edge margin, width of live area, height of live area, weight of the book, thickness of the book including cover, and thickness of the book without cover;

3. Twenty-two physical and additional characteristics of the books: total number of pages, number of pages of the table of contents, levels of contents, single work/collection, existence of Japanese-style colophon [okuzuke], existence of frontispiece, existence of foldout pages, colour printing, number of pictures, number of photographs, number of tables, number of graphs, existence of bibliography, position of bibliography, existence of preface, existence of postscript, existence of emphasis, marking of emphasis, number of indexes, numbers of index pages, names of indexes, and vertical lining/horizontal lining of the indexes;
4. Nine characteristics related to symbolic arrangements in a page: vertical lining/horizontal lining of content pages, number of columns in a page, number of lines in a page, number of characters per line, existence of page number, position of page number, existence of headers, frequency of headers, and position of headers in content page;
5. Four arrangements of sections: depth of segments, notation of caption, caption numbering with dotted decimal notation, and density of segments.

This paper mainly analyses the characteristics that belong to the second and fourth categories, since these characteristics contribute to clarifying the mode of physico-symbolic interaction between books and readers. The units of books, chapters, sections, paragraphs, and sentences are logical units that reflect information and/or knowledge. A page of a book or a double page spread, on the other hand, constitutes a physico-symbolic unit in the process of reading, which does not completely correspond to a unit of information and/or knowledge. Nevertheless, human reading is undertaken based on page spread, and there is often a clear correlation between page size and content. Thus, describing the physico-symbolic aspects of books is essential not only for revealing the form and condition of printed books as a knowledge infrastructure, but also for examining the arrangement of knowledge—which binds our understanding—in the modern era.

Methods of Measurement

Size is measured in millimetres (hereafter denoted by “mm”). Book thickness is measured in increments of 0.1 mm using a digital vernier caliper. The arrangement of pages is measured at around page 100, choosing one page that does not contain figures or tables.

We examined all the 302 sample titles in the university libraries. While *Shinsho* books are generally in paperback style, some of the old volumes in the libraries are bound in hardcover and the edges of the books are cut. Among the 302 titles, 58 are bound in hardcover.

GENERAL PHYSICO-SYMBOLIC CHARACTERISTICS OF IWANAMI SHINSHO

Columns and Lining

All the 302 books we examined have a single column. As for the lining style, 292 (96.7%) have vertical lining, six (2.0%) have horizontal lining, and four (1.3%) have both vertical and horizontal lining. The books that have horizontal lining are scattered from 1959 to 1990, and the books that have both vertical and horizontal lining appeared after 2000. We mainly

analysed the 292 books that have vertical lining. In counting the total number of pages and the thickness of books, however, we included those that have horizontal lining.

Table 2. Summary statistics of the height and width of pages, and the thickness of the books

	mean	maximum value	minimum value	standard deviation	coefficient of variation
height of pages (mm)	172.8	176	167	0.7	0.004
width of pages (mm)	105.0	111	100	1.3	0.013
thickness of books (mm)	9.1	17.9	6.9	1.3	0.149

Table 3. Statistics for the height and width of the live area, and margins

	mean	maximum value	minimum value	standard deviation	coefficient of variation
height of live area (mm)	131.8	142	124	2.5	0.018
width of live area (mm)	79.0	89	71	3.1	0.040
length of head margin (mm)	20.5	25	14	1.4	0.069
length of bottom margin (mm)	20.0	25	16	1.4	0.068
length of fore edge (mm)	11.7	19	6	2.2	0.187

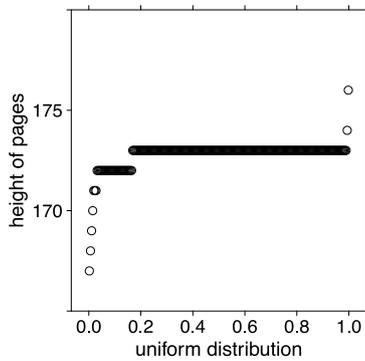


Figure 1a

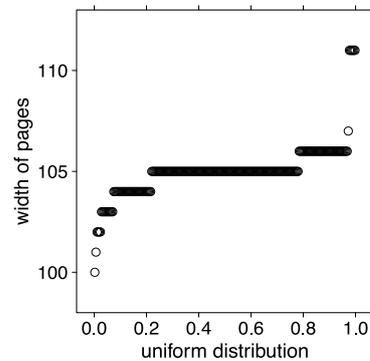


Figure 1b

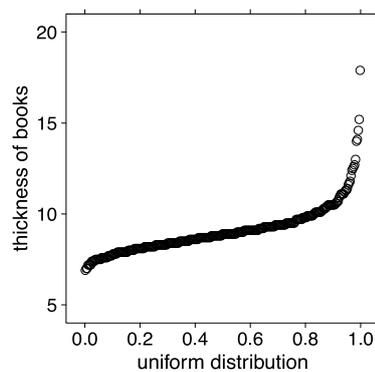


Figure 2

Page Size and Thickness of Books

Table 2 shows five summary statistics (mean, maximum value, minimum value, standard deviation, and coefficient of variation) of the height and width of pages, and the thickness of the books. In this analysis, we used the full sample of 302 books for the thickness of the books, and the 234 books not bound in hardcover and have vertical lining for the page size. We measured both the thickness of the books including cover and the thickness without cover. In this study, we analysed the latter as the thickness of the books. The table shows that there is little variation in the height and width of pages and that the thickness of books has a wider variation than the page size. The uniform-Q plots in Figures 1a and 1b show that the page heights are concentrated around 173 mm, and the page widths are clustered in the range of 104 mm to 106 mm. Figure 2 provides uniform-Q plots for the thickness of books. We observe that the book thickness has a wide variation, but are concentrated in the range 8 mm to 10 mm. Therefore, the *Iwanami Shinsho* series has some thicker books.

Size of Live Area

Table 3 shows five basic statistics for the height and width of the live area of 292 books. This table also shows basic statistics for the size of the head margin, bottom margin, and fore edge. For these three characteristics, we used 234 books that are not bound in hardcover and have vertical lining. The table shows that there is little variation in the height of the live area. We observe that unlike the page width shown in Table 2, the width of the live area has a wider variation. The uniform-Q plots in Figures 3a and 3b show that while the heights of the live area are concentrated around 133 mm, and widths are concentrated in two ranges—between 75 mm and 80 mm, and between 80 mm and 85 mm.

Number of Lines and Number of Characters

Table 4 shows basic statistics for the number of lines per page and characters per line. In this analysis, we used 289 books that have vertical lining and have an unambiguous number of characters per page. The uniform-Q plot in Figure 4a shows that most of the books have 15 or 16 lines per page, and the uniform-Q plot in Figure 4b shows that more than half of the books have 42 characters per line while some other books have 40, 41 or 43 characters.

We derived the number of characters per page and per double page spread. The latter is an important aspect because it is the standard unit of knowledge presentation in printed books. Table 5 shows the five summary statistics of total number of characters per double page spread, together with the corresponding indices for the area of double page spread. We used the 289 books that have vertical lining and an unambiguous number of characters per page for the total number of characters per double page spread; and the 234 books that are not bound in hardcover and have vertical lining for the area of double page spread. Figure 5 provides the uniform-Q plot for the number of characters per double page spread. We observe that the numbers of characters per double page spread are clustered in two groups of around 1,250 and around 1,350 characters.

Theoretical Size of Content

By multiplying the number of characters per page by the total number of pages per title, we can estimate the theoretical size of the content in terms of the number of characters. Table 6 shows the five summary statistics of the number of pages per title and the theoretical size of content. In this analysis, we used the 302 books for the total number of pages per title, and the

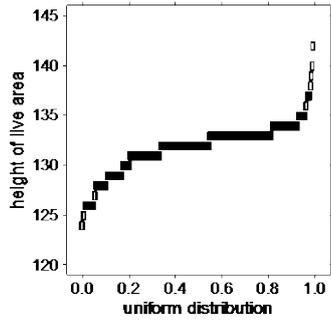


Figure 3a

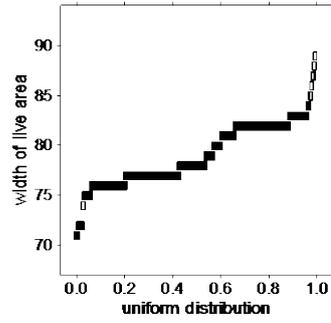


Figure 3b

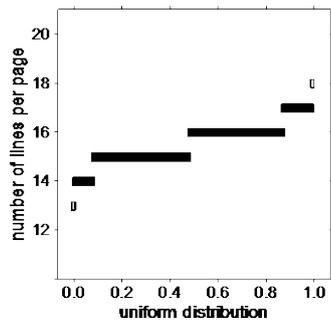


Figure 4a

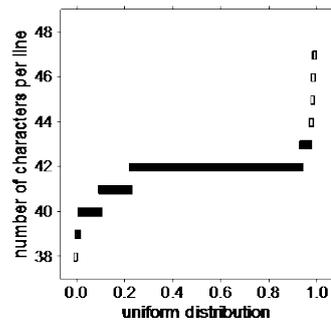


Figure 4b

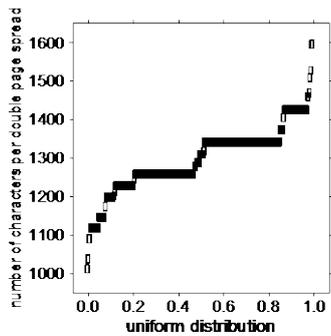


Figure 5

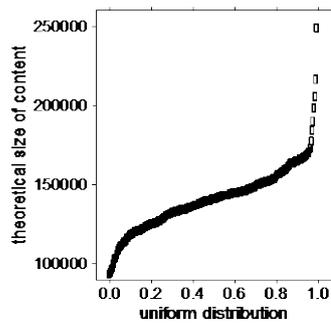


Figure 6

289 books with vertical lining and an unambiguous number of characters per page for the number of characters per page. Figure 6 provides the uniform-Q plot for the theoretical size. The size of the content varies somewhat for books of over 170,000 characters, while most of the books contain around 150,000 characters.

Brief Summary

All the sample 302 books have single column, while 292 have vertical lining. The average size of an *Iwanami Shinsho* title is 172.8 mm × 105.0 mm, and the average thickness is 9.1 mm. The average size of the live area is 131.8 mm × 79.0 mm. There are more variations in the width of live area than in the height. The average numbers of lines and characters are 15.6 lines per page and 41.8 characters per line. The number of characters per double page spread is 1,299, and the estimated average theoretical size of the content is 141,400 characters.

CHRONOLOGICAL OBSERVATIONS

We turn now to the chronological changes in the book characteristics reported in the previous section. In this analysis, we grouped the books by decade. For example, books published from 1950 to 1959 are identified as published in the 1950s. Therefore, in the analysis below, we analysed the chronological changes over nine decades from the 1930s to the 2010s. We also investigated the differences among the four colour versions that we confirmed in the previous section.

Page Size and Thickness of Books

Figure 7a plots the page heights of books chronologically, compared to a linear regression line. Figure 7b shows the moving average of the page heights. Figure 8a plots the page widths

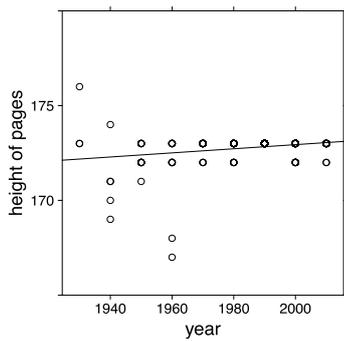


Figure 7a

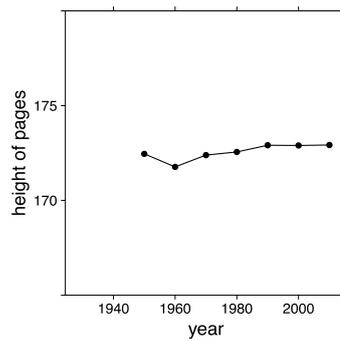


Figure 7b

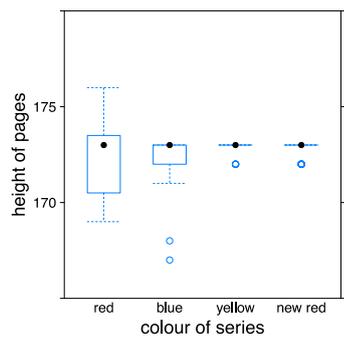


Figure 7c

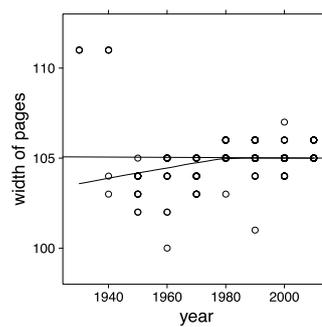


Figure 8a

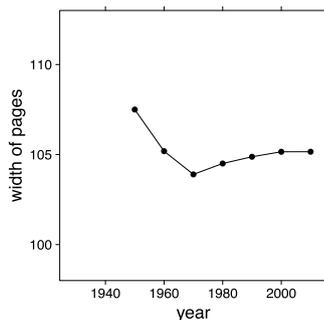


Figure 8b

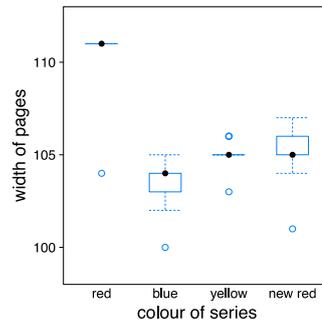


Figure 8c

chronologically, compared to a linear regression line and a spline curve (present in all similar Figures below). Figure 8b shows the same information as Figure 7b for the width. Figure 7a indicates that the page heights in the 1930s were a bit larger than that of later years and decreased in the 1940s. Since the 1950s, the values have stabilised around 172 mm or 173 mm. In Figure 8b, we observe that the page width grew larger by 5 mm from the 1930s to the 1970s, and decreased to some degree since the 1970s. Figure 7c provides the boxplots of the page heights for each colour version. Figure 8c shows the same information for the page widths. The medians of the page heights are almost the same, and the variability became smaller as the colour versions change. As for the width, only the red version has wider pages than the other versions, and it grew slightly larger from the blue version to the new red version.

Figures 9a and 9b show the chronological transitions of the book thickness and the moving average, respectively. They indicate that the book thickness was around 12 mm in the 1930s and decreased until the 1960s, then remained around 9 mm. Figure 9c shows the

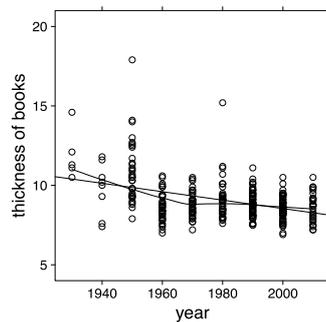


Figure 9a

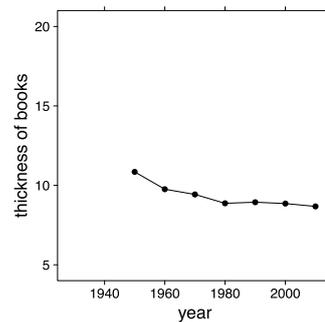


Figure 9b

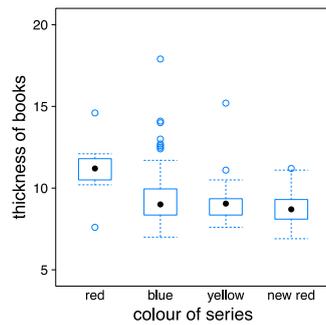


Figure 9c

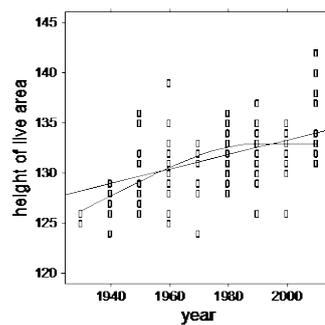


Figure 10a

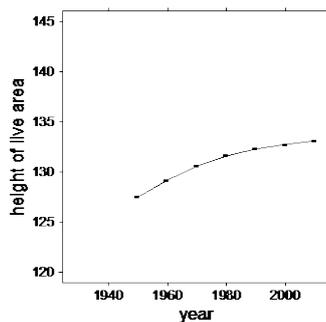


Figure 10b

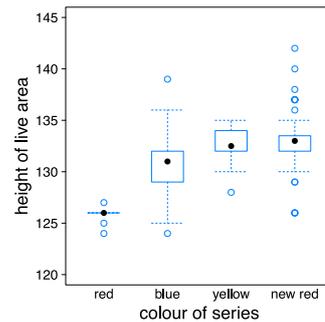


Figure 10c

boxplots of the book thickness. We observe that the red version is thicker than the other versions, and that the thickness has not changed much since the blue version.

Size of Live Area

Figures 10a and 10b show the chronological transitions of the height of the live area and the moving average, respectively. Figures 11a and 11b show the same information for the width of live area. We observe that the height of the live area has increased steadily since the 1930s and has grown 7 mm since publication started. On the other hand, the width of the live area has not changed very much, growing only 1 or 2 mm larger, if at all. The boxplots of the height of the live area (Figure 10c) indicate that the heights of the live area for the red version are concentrated around 126 mm, and those for the new red version around 133 mm. The boxplots of the width of the live area (Figure 11c) indicate that the median of the new red version is over 80 mm.

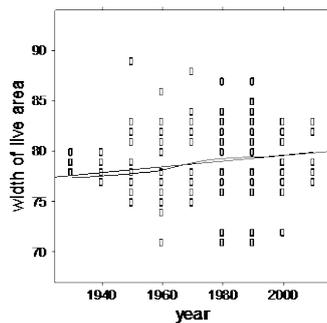


Figure 11a

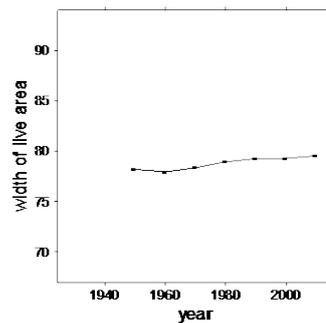


Figure 11b

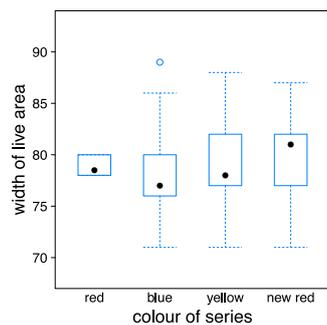


Figure 11c

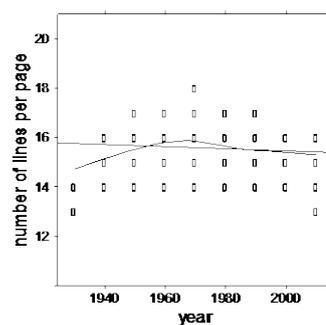


Figure 12a

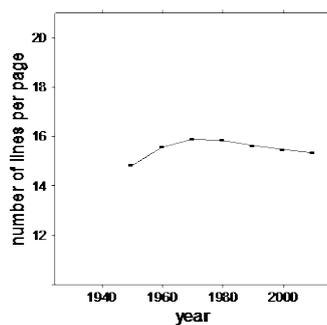


Figure 12b

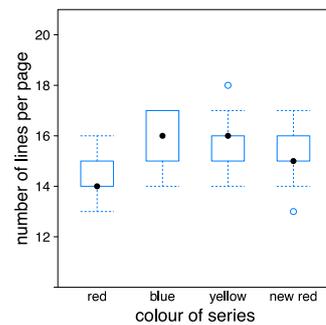


Figure 12c

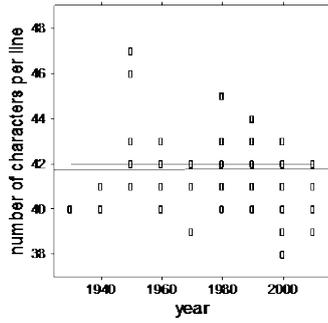


Figure 13a

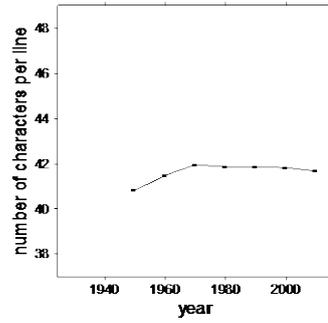


Figure 13b

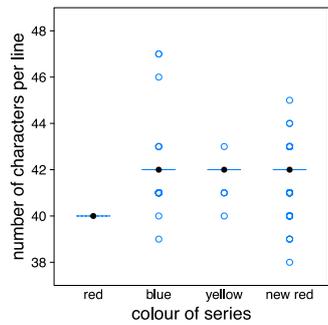


Figure 13c

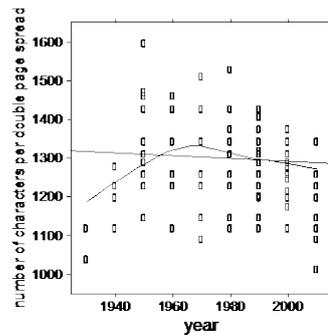


Figure 14a

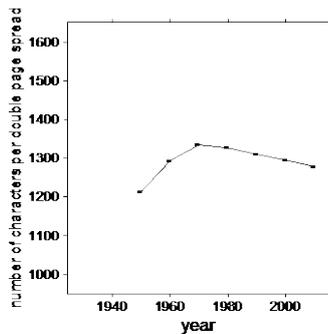


Figure 14b

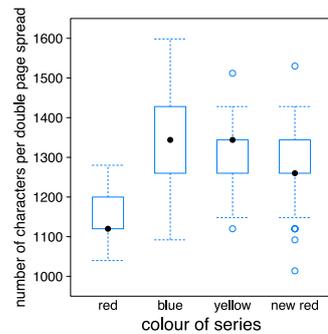


Figure 14c

Number of Lines and Number of Characters

Figures 12a and 12b show the transitions over time of the number of lines per page and the moving average, respectively. Figures 13a and 13b show the same information for the number of characters per line. The number of lines per page hits its peak at around 16 in the 1960s and 1970s. Since the 1970s, it has decreased by about one line. In Figure 12a, we observe that from the 1950s to the 1990s there are some books that have 17 lines per page, but there are books that have less than 16 lines per page after the 2000s. The number of characters per line remains stable around 42. However, Figure 13a indicates that it is under 41 before the 1940s. Figure 12c shows the boxplots of the number of lines per page, and we observe that the median of the red version is 14, and that of the blue and yellow versions is 16. We also observe that while the variations of the yellow and new red versions are almost the same, the medians of the yellow version and the new red version are different. Figure 13c shows the

boxplots of the number of characters per line, and we observe that it is only the red version whose median is 40, with a median of 42 for the other versions.

Figures 14a and 14b provide the chronological transitions of the number of characters per double page spread and the moving average, respectively. The number of characters per double page spread increased by 150 from the 1930s to the 1960s, and hit its peak in the 1960s, then decreased since the 1970s. The boxplots of the number of characters per double page spread (Figure 14c) indicate that the blue version has wider variation than the others.

Theoretical Size of Content

Figures 15a and 15b provide the chronological transitions of the number of pages per title and their moving average, respectively. We observe that the number of pages per title increased by a small amount. The boxplots of the number of pages (Figure 15c) indicate a tendency for the number of pages to increase a small amount as the colour version changes. The theoretical

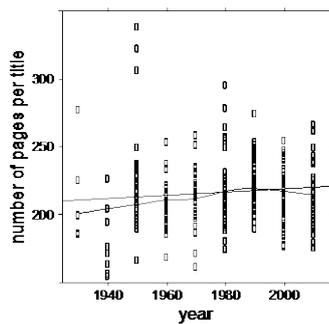


Figure 15a

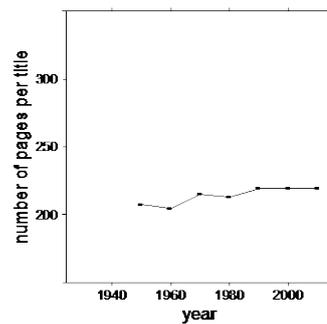


Figure 15b

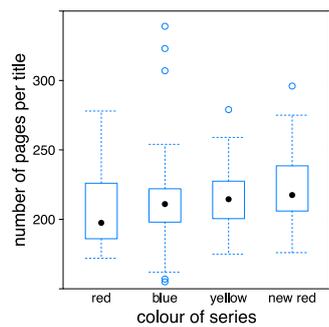


Figure 15c

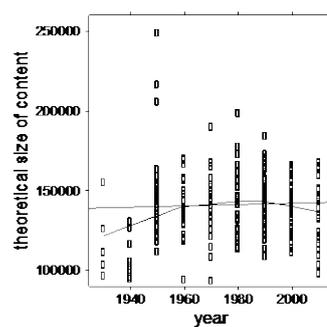


Figure 16a

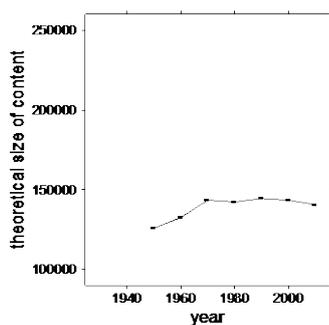


Figure 16b

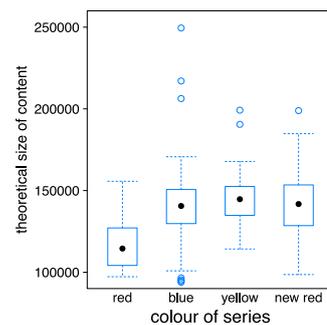


Figure 16c

size of content and the moving averages are plotted chronologically in Figures 16a and 16b. The theoretical size increased by 2,000 characters from the 1930s to the 1970s, and then flattened out around 140,000 characters until the 1990s. From the 1990s, however, it has barely decreased. Figure 16c shows the boxplots of the theoretical size of content, and we observe that except for the red version, the values of each version cluster around the same interval.

CONCLUSION AND OUTLOOK

We have analysed the physical and physico-symbolic characteristics of the pages and live area of *Iwanami Shinsho*. This study revealed the following interesting points. First, there is wider variation for the width of the live area than the width of pages. Second, the number of characters per double page spread is concentrated in two groups, around 1,250 and 1,350 characters. Third, the height of the live area has increased by 7 mm over the past 70 years, while the page height has not changed much. Fourth, the number of characters per double page spread hit its peak in the 1960s and has decreased since the 1970s. Lastly, the theoretical size of content is stable around 140,000 characters from the 1950s to the 1990s but has tended to decrease from the 1990s onward.

As for the differences between colour versions, the red version is different from the other versions in terms of both physical size and arrangement of characters. One of the reasons for this is that the editorial policy of *Iwanami Shinsho* changed after WWII. Since the start of the blue version, the physical and physico-symbolic aspects of *Iwanami Shinsho* series seem not to have changed at first glance. Though the physical size has not actually changed so much, the size of the live area has become vertically long since the advent of *Iwanami Shinsho*. Furthermore, the number of characters on a double page hit its peak in the blue version period and has decreased since the middle of that period.

Therefore, if we compare an *Iwanami Shinsho* book published in the 1960s with one published in the 2000s, they would not be different in physical form, but the live area of the latter would be taller than that of the former. The number of characters in their content would be almost the same, but the number of characters laid out on a double page spread in the book published in the 2000s would be lower than that of the book published in the 1960s.

The results of this descriptive study are important because our understanding is affected by the form and arrangement of knowledge representation. Moreover, the results of the study constitute basic and important points of departure for more experimental and pragmatic studies related to *Shinsho* and other book forms, such as the design of more readable arrangement of lines and characters on a page.

We plan to extend the descriptive study by comparing the results of this study with the results of other studies focusing on different classes of books. We have already investigated Japanese bestsellers in the post-WWII era and analysed the same characteristics in them. We are also currently investigating university textbooks in Japan.

ACKNOWLEDGEMENTS

This study is supported in part by JSPS Grant-in-Aid (B) JP25280120 “An Empirical Study of Changes in Print Media as Knowledge Infrastructure.”

REFERENCES

- Cave, R. & Ayad, S. (2014). *A history of the book in 100 books*. Buffalo, New York: Firefly Books.
- Eco, U. & Carrière, J. C. (2010). *N'espérez pas vous débarrasser des livres*. (T. Kudou, Trans.). Tokyo: Hankyu Communications Co., Ltd. (Original work published 2009).
- Eisenstein, E. L. (1980). *The printing press as an agent of change*. Cambridge, England: Cambridge University Press.
- Imamura, S. (2009). Shinshobon ni okeru zouhon oyobi shudai nohenyou: Iwanami Shinsho wo rei ni [The change of bookmaking and subjects of small-sized paperbacks]. *Memoirs of Taisho University*, 94, 243-253. [In Japanese]
- Imamura, S. (2014). Shinshobon no shudai han'i. [The coverage of topics of Shinsho], *Memoirs of Taisho University*, 99, 313-326. [In Japanese]
- Lyons, M. (2011). *Books: A living history*. London: Thames & Hudson.
- Merkoski, J. (2013). *Burning the page: The ebook revolution and the future of reading*. Naperville, Illinois: Sourcebooks.
- Pearson, D. (2013). *Books as history: The importance of books beyond their texts* (2nd rev. ed.). London: The British Library Publishing Division.
- The Editorial Department of Iwanami Shoten Publishers (Ed.) (1988). *Iwanami Shinsho no 50 nen* [50 years of Iwanami Shinsho]. Tokyo: Iwanami Shoten Publishers. [In Japanese]
- The Editorial Department of Yearbook of Publication (Ed.) (2015). *Shuppan Nenkan*. [Yearbook of publication]. Tokyo: Shuppan News Co., Ltd. [In Japanese]
- Yamada, S., Shimura, R., Umino, B., Toda, S. & Kageura, K. (2015). Physico-symbolic characteristics of Japanese bestsellers since WWII. Paper presented at the 6th *International Conference on Asia-Pacific Library and Information Education and Practice*, 28-30 October 2014, Manila, Philippines.