

Sensitivity of Indus Script to Site and Type of Object*

Nisha Yadav

Tata Institute of Fundamental Research, India

The Indus valley civilisation (2600 to 1900 BCE) left behind its undeciphered script on several small, inscribed objects. In our earlier work, we have investigated the structural aspects of the Indus script, the design of its signs and the nature of the inscribed patterns on these objects. Here we study these inscribed objects with respect to their type and the sites where they were found. The distribution of the text lengths and the usage of signs on these objects indicate subtle variations in the context of their site of occurrence and type. Some signs have affinity to specific sites or types of objects. By comparing the text length distributions and the usage of signs on these inscribed objects, we cluster the sites of occurrence and the types of objects based on their level of similarity. The study provides an understanding of the sensitivity of the Indus script to different sites and types of objects.

Keywords: Undeciphered scripts, ancient scripts, Indus valley civilisation, inscribed objects, Harappan script, Indus script - site of occurrence and medium of writing, clustering of sites and types of objects

1. Introduction

The Indus valley civilisation flourished in the northwestern part of the Indian subcontinent from 2600 to 1900 BCE though its roots can be traced

* I wish to acknowledge the support of Jamsetji Tata Trust for this work. I also wish to thank Dr. Iravatham Mahadevan for his insightful guidance. I am thankful to Prof. M. N. Vahia and Dr. Ambuja Salgaonkar for their critical suggestions and inputs. I would like to thank Mr. Hrishikesh Joglekar for his help in designing the software used for generating the text images. I thank Prof. J. M. Kenoyer and Harappa.com for their kind permission to use the image of an Indus seal in the paper and for their continuing support.

SCRIPTA, Volume 5 (September 2013): 67-103

© 2013 The Hunmin jeongeum Society

to early sites such as Mehrgarh around 7000 BCE. At its peak it was spread over an area of about a million square kilometres in several urban, semi-urban and rural settlements and was the largest Bronze Age civilisations in the world (for review see Wright 2010, Agrawal 2009, 2007, Possehl 2002, Kenoyer 1998). The entire evidence about the civilisation is archaeological and its relation to later cultures in the subcontinent is unclear.

One of the intriguing aspects of the Indus valley civilisation is its script that has not been deciphered. This is due to the extreme brevity of the texts, absence of bilingual or multilingual texts, lack of knowledge about their language(s) and apparent discontinuity in the traditions at the decline of the Indus valley civilisation. Possehl (1996) provides an excellent critical review of some of the various attempts to decipher the script (see also Parpola 2005, Mahadevan 2002). In spite of these efforts, the problem of the Indus script lies unresolved with no universal consensus on any of the proposed interpretations.

The texts constituting the corpus of Indus script are predominantly found on objects such as seals and sealings (generally made of steatite or terracotta). Most often, these objects are a few square centimetres in size. An example of a square Indus seal is shown in Fig. 1a. It has an Indus text consisting of eight signs at the top, a field symbol (Unicorn) at the centre

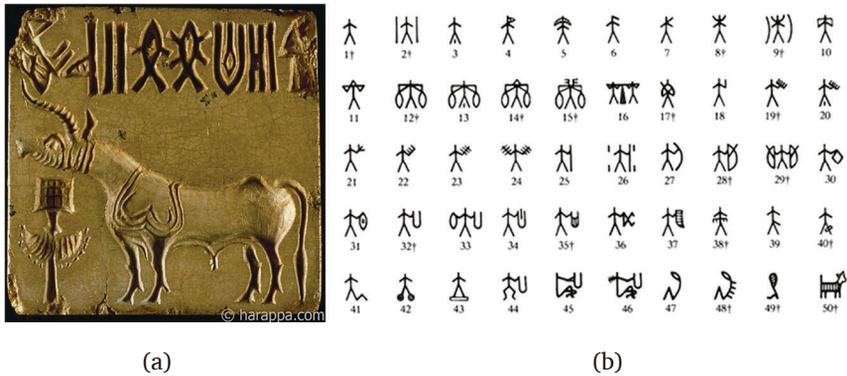


Figure 1. (a) A square seal with an Indus text on top, a field symbol (Unicorn) in the middle and a decorated object on the bottom left (Copyright Harappa Archaeological Research Project/J. M. Kenoyer, Harappa.com, Courtesy Dept. of Archaeology and Museums, Govt. of Pakistan). The seal is about 5 cm x 5 cm in size. (b) Sign numbers 1 to 50 from the sign list of Mahadevan's concordance of the Indus script (Mahadevan 1977).

and a decorated object near its face. A small subset of the sign list of the Indus script (sign numbers 1 to 50) from Mahadevan's concordance of the Indus script (Mahadevan 1977) is shown in Fig. 1b.

Around 3700 such objects have been unearthed from several Indus sites situated in India, Pakistan and some West Asian sites. They are catalogued in the three volumes of the Corpus of Indus Seals and Inscriptions (CISI 1-3 by Joshi and Parpola (1987), Shah and Parpola (1991) and Parpola et al. 2010 respectively). These volumes provide a photographic documentation of almost all of the inscribed objects discovered from various Indus sites. Images of some of these objects are also available at www.harappa.com. The texts in the Indus script from these inscribed objects are compiled and catalogued in the three concordances of the Indus script created by Mahadevan (1977), Koskenniemi and Parpola (1982, see also Parpola 1994) and Wells (2006). The first two concordances are largely in agreement with each other though small differences in the identification of a few signs remain.

We use the electronic concordance of Mahadevan (1977, henceforth referred to as M77) which records 3548 lines of texts from 2906 inscribed objects along with their contextual information such as the type of object, site of occurrence, field symbols etc. The sign list of M77 consists of 417 signs¹. While there is a general increase in the design complexity of the signs with the sign number in M77 (within groups of similar looking signs), completely different signs also appear consecutively in the sign list (see for example, sign numbers 46 and 47 in Fig. 1b). Hence, the proximity of sign numbers does not necessarily indicate similarity in design or any other relation between signs and each sign is considered as an independent entity in M77.

M77 was created in 1977 incorporating about 80% of the inscribed objects available today. Another concordance of comparable size by Koskenniemi and Parpola (1982) also belongs to the same period. M77 includes the Indus texts from the inscribed objects catalogued in CISI volumes 1 and 2 (Joshi and Parpola (1987), Shah and Parpola (1991)) as

¹ The serial number of the signs used in this paper is as in the sign list of M77. This is a normalised sign list and at times, a sign has one or more variants that are listed in Appendix 1 of M77 (Mahadevan: 785-792). As a convention followed in the present paper, the texts depicted as strings of sign images are to be read from right to left, whereas the texts represented as strings of sign numbers are to be read from left to right.

well as other objects from older excavations at Mohenjodaro and Harappa that have been published recently in the third photographic volume of the Corpus of Indus Seals and Inscriptions (CISI-3: Parpola et al. 2010). CISI-3 also includes about 1000 inscribed objects from the Harappa Archaeological Research Project (HARP) excavations (1986-2007). The HARP data in CISI-3 includes inscribed objects with or without script and some of them are also damaged or broken. We estimate that about 500 objects with unambiguous texts from the HARP data are not included in M77. The additional data has added only a few signs to the sign list of the Indus script and most of these objects belong to the same categories as defined in M77. This amounts to about 15% of the data available till date and introduces corresponding uncertainty in the conclusions. However, the present study is indicative of the broad trends in the Indus writing. Inclusion of the new data will further improve the statistical significance of the results presented here. Within these limitations, the present analysis provides interesting insights into the written material of the Indus valley civilisation.

In our earlier work we have studied various structural aspects of the Indus script (Yadav 2012, Yadav et al. 2010, Rao et al. 2010, Rao et al. 2009a, b and Yadav et al. 2008a, b), its sign design (Yadav and Vahia 2011b) and different types of patterns on these inscribed objects (Yadav and Vahia 2011a, Sinha et al. 2011, Vahia and Yadav 2010). In this paper, we undertake an analysis of the Indus script with respect to its site of occurrence and type of object.

2. Framework for the Study of the Indus Script

The Indus texts are found on various types of objects excavated from several sites of the Indus valley civilisation and most often they are associated with different kinds of patterns (Yadav and Vahia 2011a). A comprehensive framework to understand the context and usage of the Indus script is given in Fig. 2. It has two main components that deal with the following aspects

- 1) structural studies of the Indus script (shown at the centre), and
- 2) contextual studies of the Indus script (given in the outer loop).

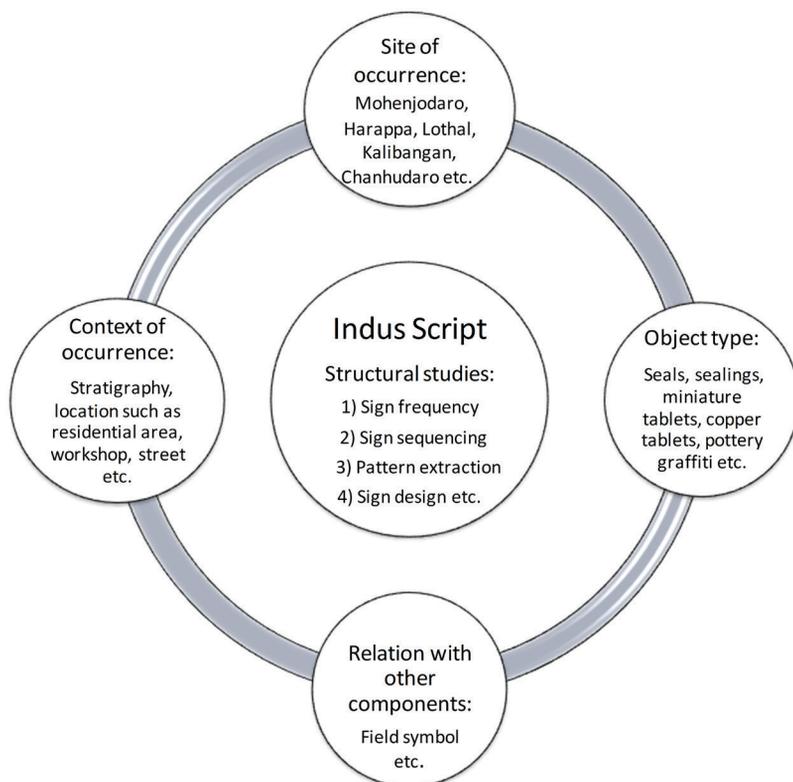
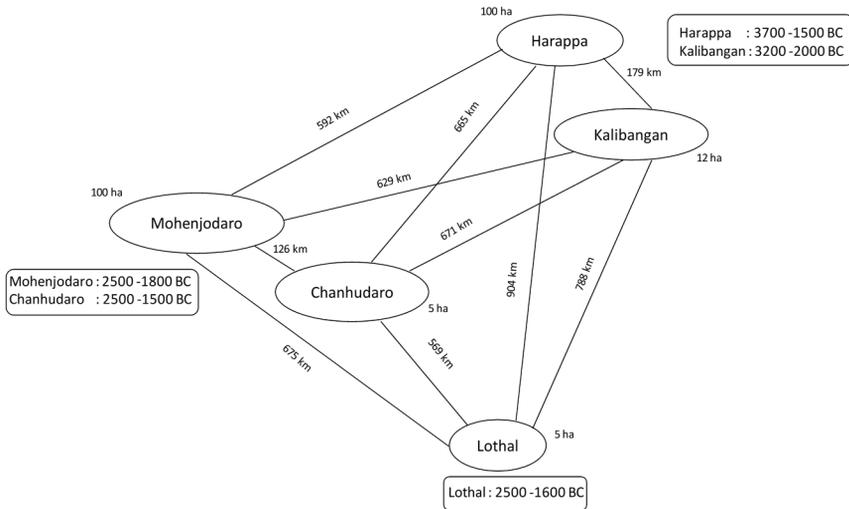


Figure 2. A comprehensive framework to understand the context and usage of the Indus script. The central component deals with the structural aspects of the Indus script and components related to the contextual aspect of the Indus script are shown in the outer loop.

The structural aspects of the Indus script such as the sequencing of signs in the Indus texts were addressed in some of our earlier work on the Indus script (Yadav 2012, Yadav et al. 2010, Rao et al. 2010, Rao et al. 2009a, b, Yadav et al. 2008a, b) and the design of Indus signs was investigated in Yadav and Vahia (2011b). The contextual aspects of the Indus script include

- 1) site of occurrence of the object,
- 2) type of object,
- 3) relation of the script with other inscribed components on the object, and,



Note: The lengths of the connecting lines are for representative purpose only.

Figure 3. Relative location and size of the Indus sites which account for about 98% of the data in M77. Location of the sites in the figure is relative to their geographical locations and the distance between two sites is taken as the shortest possible distance between them (Data of location, period and area of sites is taken from Possehl 1999).

- 4) other information including stratigraphy, precise location of the object etc.

We briefly discuss each of these contextual components below.

2.1 Site of Occurrence

M77 includes data from 19 Harappan sites including Mohenjodaro (MD), Harappa (HP), Lothal (LL), Kalibangan (KB), Chanhudaro (CH), Other Harappan sites (OH) and five West Asian sites (WA). Fig. 3 gives the relative locations, distance and extent of the five Indus sites that contribute to about 98% of the data in M77.

The sites in Fig. 3 form three clusters by proximity, namely Mohenjodaro-Chanhudaro, Harappa-Kalibangan and Lothal. The statistics of the number of objects, lines of texts and sign occurrences at each of these sites is given in Table 1.

Table 1. Distribution of objects, texts and sign occurrences at different sites in M77.

	Site of occurrence						
	MD	HP	LL	KB	CH	OH	WA
Number of objects (Percentage of total)	1540 (52.99)	985 (33.90)	165 (5.68)	99 (3.41)	66 (2.27)	34 (1.17)	17 (0.58)
Number of lines of texts (Percentage of total)	1716 (48.37)	1375 (38.75)	196 (5.52)	117 (3.30)	86 (2.42)	40 (1.13)	18 (0.51)
Number of sign occurrences (Percentage of total)	7281 (54.45)	4359 (32.60)	760 (5.68)	423 (3.16)	327 (2.45)	141 (1.05)	81 (0.61)

* Each line of text is considered as a separate entity in multi-lined texts.

As can be seen from Table 1, Mohenjodaro, Harappa, Lothal, Kalibangan and Chanhudaro together account for about 98% of the data in M77. The total percentage contribution of the remaining sites (Other Harappan sites and West Asian sites) is less than 2% in M77.

2.2 Type of Object

M77 includes data from various types of inscribed objects such as seals (S), sealings (SL), miniature tablets (MT), copper tablets (CT), pottery graffiti (PG), ivory or bone rods (IB), bronze implements (BI) and other miscellaneous objects (Misc.). We use these terminologies directly from M77 and they are defined as follows:

- 1) seals are objects with text deeply engraved on them in *intaglio* in the reverse direction,
- 2) sealings are positive impressions in relief by seals or special moulds on impressible material like clay, faience or metal,
- 3) miniature tablets are like sealings but are of exceptionally small size,
- 4) copper tablets and bronze implements are objects made of copper and bronze respectively,
- 5) pottery graffiti includes texts on pottery,
- 6) ivory and bone rods include objects made of animal bones, and,
- 7) miscellaneous objects include stone objects, bracelets, dice, carnelian tablets, terracotta ball, brick etc. These objects do not fall in any of the

Table 2. Distribution of texts and sign occurrences on different types of objects in M77.

	Type of object							
	S	SL	MT	CT	PG	IB	BI	Misc.
Number of objects (Percentage of total)	1814 (62.42)	511 (17.58)	272 (9.36)	135 (4.65)	119 (4.09)	29 (1.00)	11 (0.38)	15 (0.52)
Number of lines of texts (Percentage of total)	1963 (55.33)	731 (20.60)	499 (14.06)	162 (4.57)	127 (3.58)	29 (0.82)	17 (0.48)	20 (0.56)
Number of sign occurrences (Percentage of total)	8312 (62.16)	2582 (19.31)	1261 (9.43)	718 (5.37)	281 (2.10)	104 (0.78)	64 (0.48)	50 (0.37)

* Each line of text is considered as a separate entity in multi-lined texts.

above categories.

In M77, the objects classified as sealings may include potentially more than one category of objects with possible differences in distribution. For example, seal impressions on clay tags, seal impressions on pots and moulded tablets may belong to different categories. It is possible that division of sealings into more categories can provide further insights. However, as we shall see later, even the present classification of objects shows that sealings (probably of all the three classes) are predominantly found at one site. The total number of sealings in M77 is 511 and further division of this class may reduce the sample size to limits of error.

The statistics of the number of lines of texts and sign occurrences on different types of objects is given in Table 2. As can be seen from Table 2, seals, sealings, miniature tablets, copper tablets and pottery graffiti together account for about 98% of the data in M77. The total percentage contribution of the remaining types of objects (ivory or bone rods, bronze implements and miscellaneous objects) in M77 is less than 2%. Although the objects in some of these categories can be further classified on the basis of their shapes (or type of boss in case of seals) and other features, these finer details are not considered in the present study.

2.3 Relation of the Indus Script with other Components on Inscribed objects

The inscribed objects often contain different kinds of patterns in addition to the Indus script. M77 refers to these patterns as field symbols and broadly classifies them into the following types:

- 1) animals,
- 2) reptiles, fish, bird etc.,
- 3) trees and leaves,
- 4) anthropomorphic forms,
- 5) scenes with anthropomorphic and animal figures, trees and other objects,
- 6) various symbols, motifs and geometrical patterns, and,
- 7) damaged or illegible field symbol.

Yadav and Vahia (2011a) have classified various components of the inscribed surfaces into twelve categories (such as the Indus script, animal motifs, manger or decorated object, composite animals, multi-headed animals etc.) and analysed the percentage contribution of each of these components in the corpus of inscriptions. Based on the analysis, the Indus script is the most frequent component on the inscribed surfaces (67%), followed by animal motifs (35%) and the decorated object or manger generally near the face of an animal (22%). The geometric and symmetric patterns on the inscribed objects are discussed in detail in Vahia and Yadav (2010).

2.4 Other Contextual Information

Other contextual information includes aspects such as stratigraphy and precise location (residential area, workshop, street etc.) of the object at a site. Since the stratigraphy of most of the data remains unclear, we do not discuss it here.

The relation of the Indus script with its contextual aspects along with its structural studies provides a framework for understanding the role of the Indus script in the lives of the Indus people. In the present study, we initiate the analysis of the Indus script with respect to two aspects of the inscribed objects, namely, their site of occurrence and their type.

3. Study of Indus Script with respect to Site and Type of object

An important aspect of the relevance of the Indus script to the Indus valley civilisation is the study of the similarities and differences in the written material found at various Indus sites and on different types of objects. In the present study, we analyse the data along the following lines

- 1) distribution of inscribed objects with respect to sites and types of objects,
- 2) text length distribution with respect to sites and types of objects,
- 3) sign usage at different sites, and,
- 4) sign usage on different types of objects.

3.1 Distribution of Inscribed Objects with respect to Sites and Types of Objects

In the present section, we discuss the relative distribution of the inscribed material at different sites. While some of these features may be an artifact of sampling or the excavation procedure, the results are indicative of broad trends. The percentage contribution of various types of objects at each site is shown in Fig. 4. Following conclusions can be derived from Fig. 4:

- 1) Most of the inscribed material found in Mohenjodaro is in the form of seals followed by copper tablets and sealings.
- 2) Harappa has comparable percentage of seals, sealings and miniature tablets. Lothal has nearly equal number of seals and sealings. A detailed analysis of the data shows that the sealings at Lothal are mostly in the form of clay tags with impressions of one or more seals embossed on them.
- 3) The inscribed data in Kalibangan is mostly in the form of seals followed by sealings and pottery graffiti.
- 4) Most of the inscribed material found in Chanhudaro is in the form of seals.
- 5) The inscribed material from Other Harappan sites is predominantly in the form of seals and pottery graffiti.
- 6) West Asian sites have a large percentage of their inscribed material as seals and a small percentage as sealings.

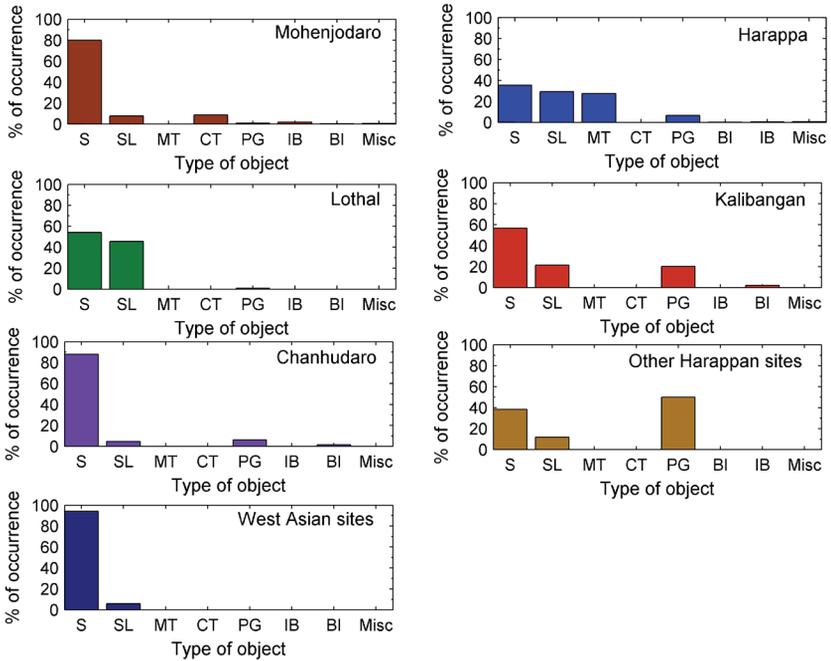


Figure 4. Percentage contribution of different types of objects at each site. The data at each site is normalised to 100% to compare the relative contribution of various types of objects across the sites.

The percentage contribution of various sites to each type of object is shown in Fig. 5. Following conclusions can be derived from Fig. 5:

- 1) Mohenjodaro accounts for highest percentage of seals followed by Harappa.
- 2) Harappa accounts for highest percentage of sealings followed by Mohenjodaro. About 23% of the texts on the sealings from Harappa are duplicates of the four most frequent texts in M77 (see section 5).
- 3) Miniature tablets are found only in Harappa and about 36% of the texts on these miniature tablets are duplicates of the four most frequent texts in M77 (see section 5).
- 4) Copper tablets and ivory or bone rods are predominantly found in Mohenjodaro.
- 5) Almost half of the pottery graffiti is found at Harappa.
- 6) Pottery graffiti and bronze implements appear at almost all major sites but in Lothal they are extremely rare.

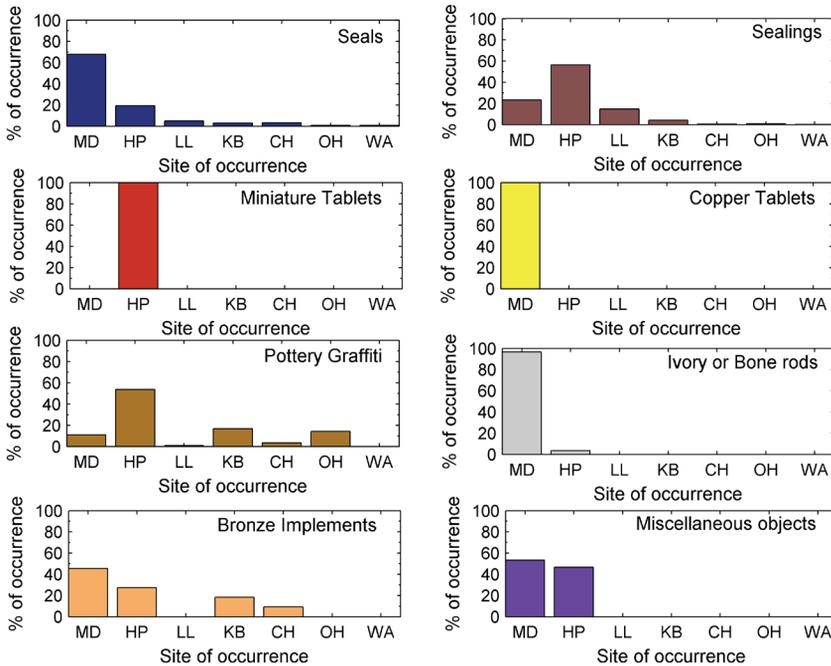


Figure 5. Percentage contribution of different sites to each type of object. The data from each type of object is normalised to 100% to compare the relative contribution of various sites to different types of objects.

7) The inscribed material from West Asian sites accounts for less than 1% of the total sample size and hence is not visible in Fig. 5.

3.2 Text Length Distribution at Different Sites and on Different Types of Objects

The length (or the total number of signs) of an Indus text in a single line varies from 1 to 14 signs in M77 (Yadav et al. 2010). The average length of the Indus texts is about 5 signs. The distribution of texts of varying lengths at different sites and object types are given in Figs. 6 and 7 respectively. Following conclusions can be derived from Figs. 6 and 7:

1) The distribution of text lengths at Mohenjodaro is broadly similar to that of seals. This is consistent with the fact that seals are most frequent at Mohenjodaro (Fig. 5).

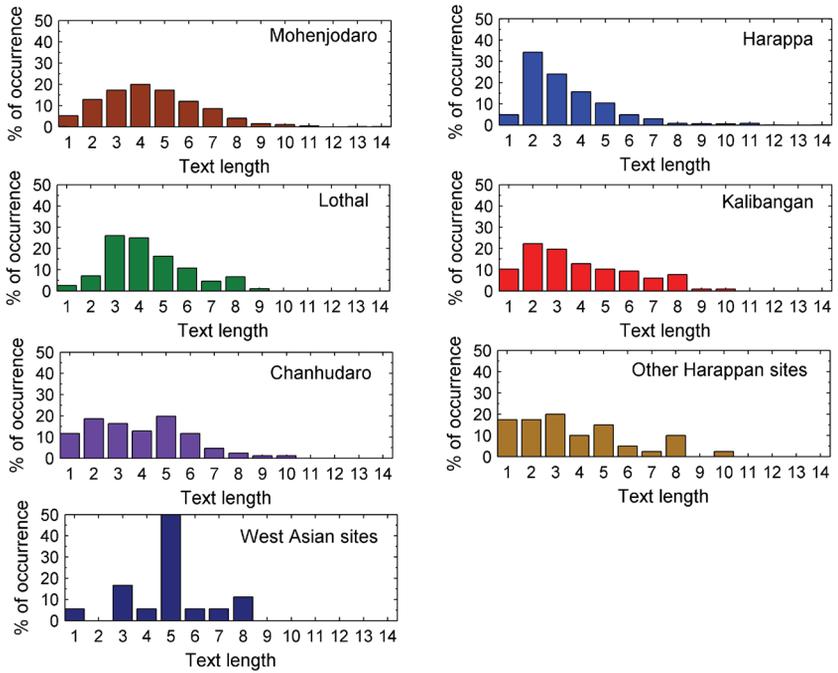


Figure 6. Percentage contribution of texts of lengths 1 to 14 at different sites. The data at each site is normalised to 100% to compare the relative contribution of texts of different lengths across the sites.

- 2) The most frequent text length at Harappa is two. This is mainly due to the presence of a large number of miniature tablets and sealings at Harappa with the frequent texts predominantly of length two (see section 5).
- 3) The distribution of text lengths at Harappa is roughly similar to that of sealings and miniature tablets. This is consistent with the fact that sealings and miniature tablets are most frequent at Harappa (Fig. 5).
- 4) In comparison to other sites, Lothal, Kalibangan, Other Harappan sites and West Asian sites have relatively larger fraction of texts of length eight.
- 5) In relative terms, West Asian sites have significantly large number of texts of length five.
- 6) It should be noted that there is a monotonic fall in the percentage of texts with increase in text length in Mohenjodaro from text length of four onwards. Similar trend is seen in Harappa from text length of two onwards.

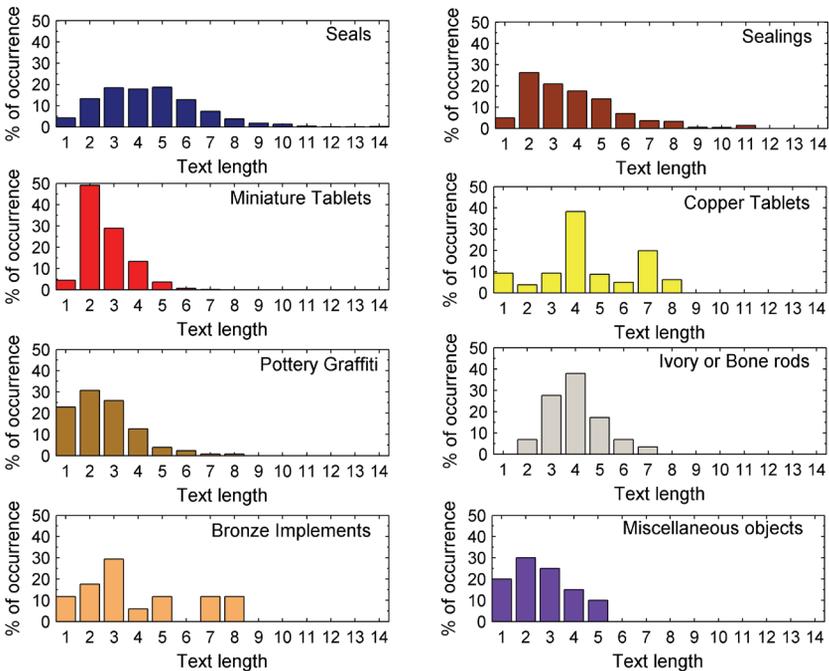


Figure 7. Percentage contribution of texts of lengths 1 to 14 on different types of objects. The data from each type of object is normalised to 100% to compare the relative contribution of texts of different lengths to various types of objects.

- 7) Texts of lengths 9 to 14 are found exclusively on seals and sealings.
- 8) Texts on miniature tablets have lengths mostly in the range of 2 to 4 signs. There are very few miniature tablets with just a single sign.
- 9) The most frequent text length on copper tablets is 4 followed by 7.
- 10) The length of most of the texts in pottery graffiti is in the range of 1 to 3 signs.

3.3 Sign Usage at Different Sites

Although the sign list of M77 has 417 signs, the most frequent 67 signs (listed in the Appendix) account for about 80% of the total sign occurrences in M77. Moreover, not all of these 417 signs occur at all sites and there are some signs in M77 that are unique to a specific site and do not occur elsewhere. The statistics of the total number of signs, most frequent 67 signs and unique signs at each site in M77 is given in Table 3. It also gives

Table 3. Sign usage at different sites in M77.

	M77 (All sites)	Site of occurrence						
		MD	HP	LL	KB	CH	OH	WA
Total number of signs (Percentage of total sign occurrences)	417 (100)	353 (54.45)	261 (32.60)	140 (5.68)	119 (3.16)	106 (2.45)	81 (1.05)	57 (0.61)
Number of signs from the 67 most frequent signs in M77 (Percentage of total sign occurrences at the site)	67 (80)	67 (78.99)	67 (84.01)	64 (79.61)	60 (79.91)	57 (77.68)	50 (75.18)	32 (65.43)
Number of signs unique to the site (Percentage of total sign occurrences at the site)	158 (2.01)	102 (2.57)	30 (1.01)	7 (0.92)	9 (2.36)	5 (4.59)	2 (1.42)	3 (4.94)

corresponding numbers for M77 for comparison.

It can be seen from Table 3 that the entire set of the 67 most frequent signs occur only at Mohenjodaro and Harappa. Not all of these 67 most frequent signs occur at other sites. However, the percentage contribution of these signs to the total sample size at each site is close to 80% except at West Asian sites where it deviates significantly (65%). The signs that are unique to specific sites have low frequency of occurrence and their percentage contribution to the total sample size at various sites does not exceed 5%. For instance, there are 102 signs in the sign list of M77 that occur only at Mohenjodaro, but as these signs are extremely rare, their total percentage contribution to the Mohenjodaro data is just 3%. The fractional usage of the 67 most frequent signs at each site is given in Fig. 8. The relative contribution of these signs fluctuates across different sites.

Several signs seem to have preferential usage at different sites (Fig. 8). The five signs from each site having highest deviation from their expected occurrences in the positive and negative directions are listed in Tables 4a and 4b respectively.

Following features are evident from Fig. 8 (a to g):

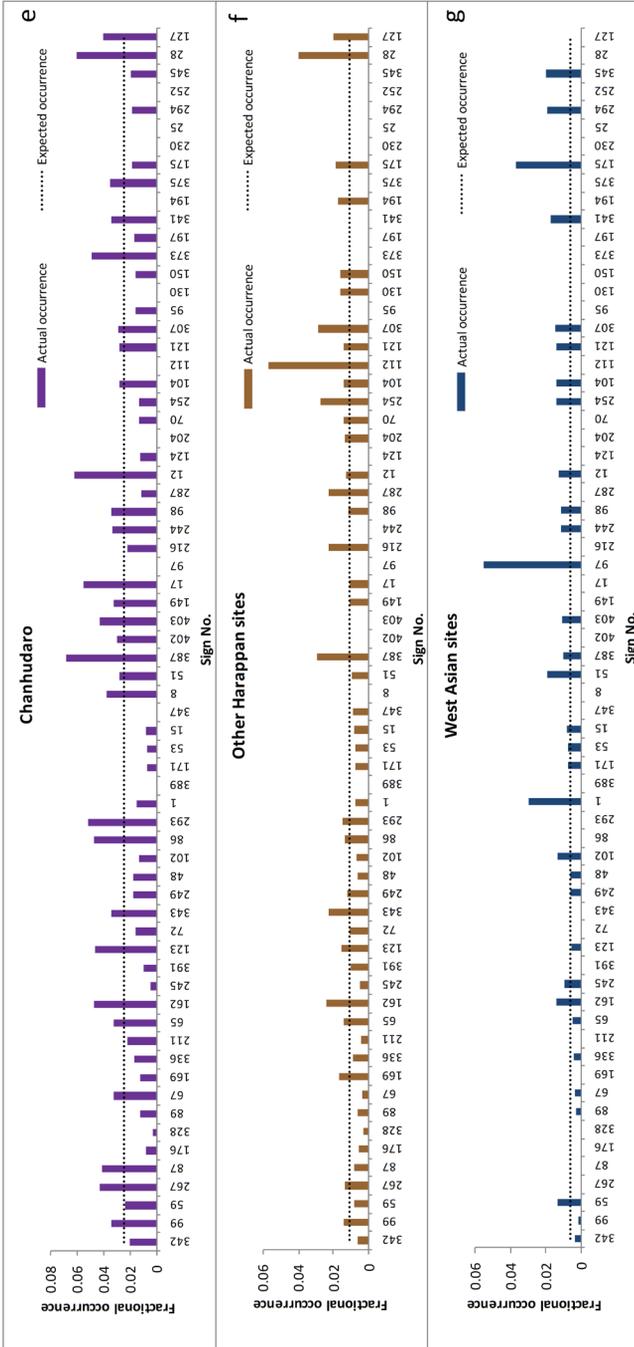


Figure 8. Fractional usage of the 67 most frequent signs at different sites (sorted in the descending order of their frequency in M77): (a) Mohenjodaro, (b) Harappa; (c) Lothal, (d) Kalibangan, (e) Chanhudaro, (f) Other Harappan sites, (g): West Asian sites (See Table in the Appendix for images and frequency of signs marked on the x-axis). The line of expected occurrence for each site is derived from its sample size in M77.

Table 4a. Signs having more than expected occurrence at various sites. The signs are listed such that the most frequent sign with highest excess is listed first.

MD	HP	LL	KB	CH	OH	WA
 25	 95	 121	 121	 387	 112	 97
 124	 328	 98	 1	 12	 28	 175
 97	 176	 254	 307	 28	 387	 1
 194	 89	 197	 175	 17	 307	 345
 245	 307	 343	 127	 293	 254	 51

Table 4b. Signs having less than expected occurrence at various sites. The signs are listed such that the most frequent sign with highest deficit is listed first.

MD	HP	LL	KB	CH	OH	WA
 95	 98	 307	 345	 389	 389	 267
 328	 121	 95	 294	 347	 8	 87
 176	 17	 345	 254	 97	 402	 176
 307	 97	 176	 197	 204	 403	 328
 121	 1	 328	 160	 112	 97	 169

- 1) The usage of sign numbers 95, 328 and 176 is relatively low at Mohenjodaro and high at Harappa. These signs are part of the four most frequent texts that duplicate several times on sealings and miniature tablets from Harappa (see section 5).
- 2) The usage of the 67 most frequent signs is relatively uneven at Lothal. Sign numbers 121, 98 and 254 are heavily used in Lothal while the sign numbers 307, 95 and 345 do not occur at all.
- 3) In Kalibangan, sign numbers 121, 1 and 307 are most frequently used.

- 4) The distribution of sign usage in Chanhudaro is non-uniform with several signs deviating significantly from their expected occurrence.
- 5) While the statistics of the signs at Other Harappan sites and West Asian sites is low, the usage of sign number 112 is unusually high at Other Harappan sites and sign numbers 97, 175 and 1 are most frequent at West Asian sites.

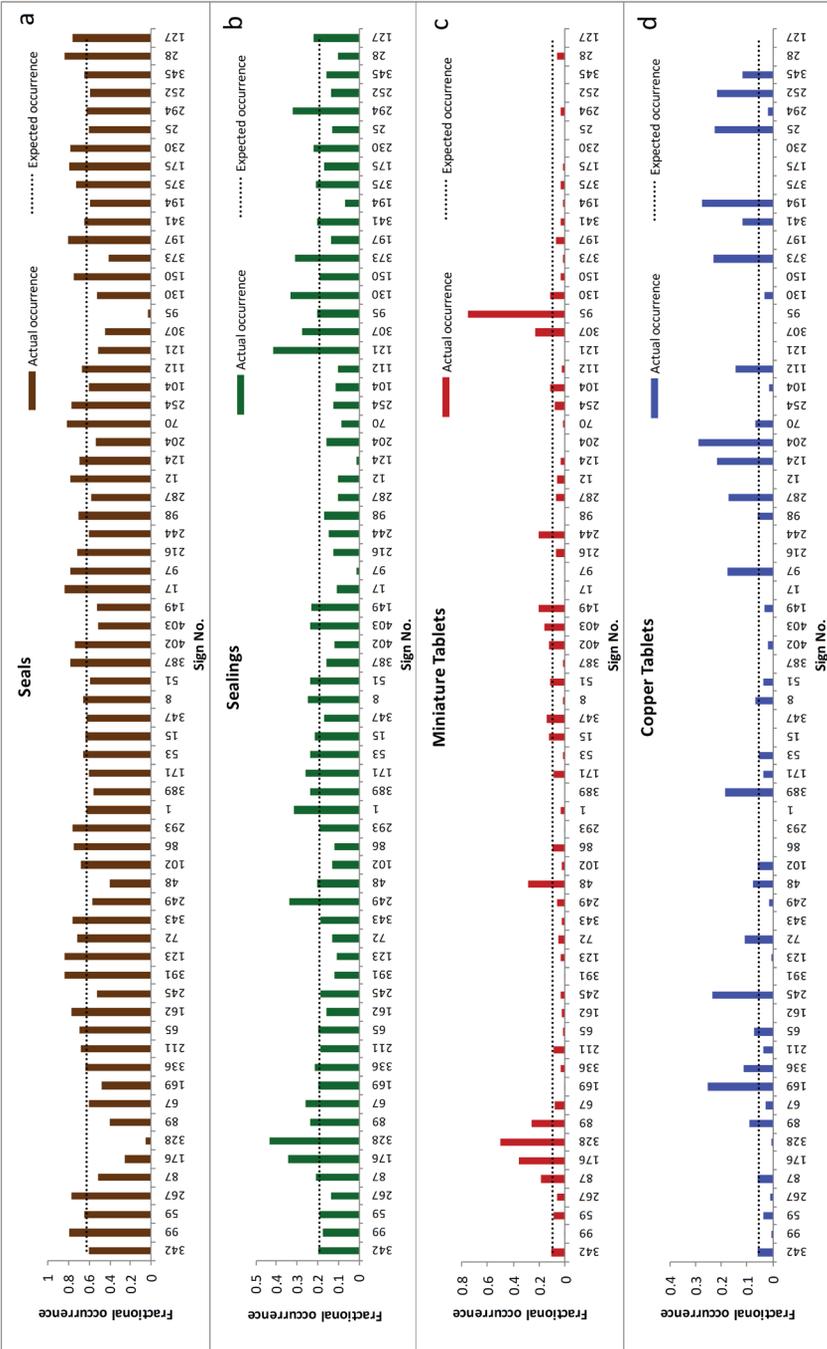
3.4. Sign Usage on Different Types of Objects

As in the case of sites, not all signs from the 417 signs in the sign list of M77 occur on all types of objects and there are some signs that are unique to a specific type of object and never occur on other types of objects. The statistics of the number of signs, the most frequent 67 signs and unique signs on various types of objects is given in Table 5. It also gives corresponding numbers for M77 for comparison.

It can be seen from Table 5 that the entire set of the 67 most

Table 5. Sign usage on different types of objects in M77.

	M77 (All objects)	Type of object							
		S	SL	MT	CT	PG	IB	BI	Misc.
Total number of signs (Percentage of total sign occurrences)	417 (100)	381 (62.16)	184 (19.31)	119 (9.43)	65 (5.37)	107 (2.10)	42 (0.78)	42 (0.48)	43 (0.37)
Number of signs from the 67 most frequent signs in M77 (Percentage of total sign occurrences on the object)	67 (80)	67 (79.49)	67 (83.08)	57 (85.25)	41 (80.50)	54 (76.51)	30 (80.77)	23 (60.94)	21 (52.00)
Number of signs unique to the type of object (Percentage of total sign occurrences on the type of object)	196 (3.78)	169 (5.52)	9 (0.81)	4 (0.56)	4 (1.11)	4 (1.42)	0 (0)	3 (4.69)	3 (6.00)



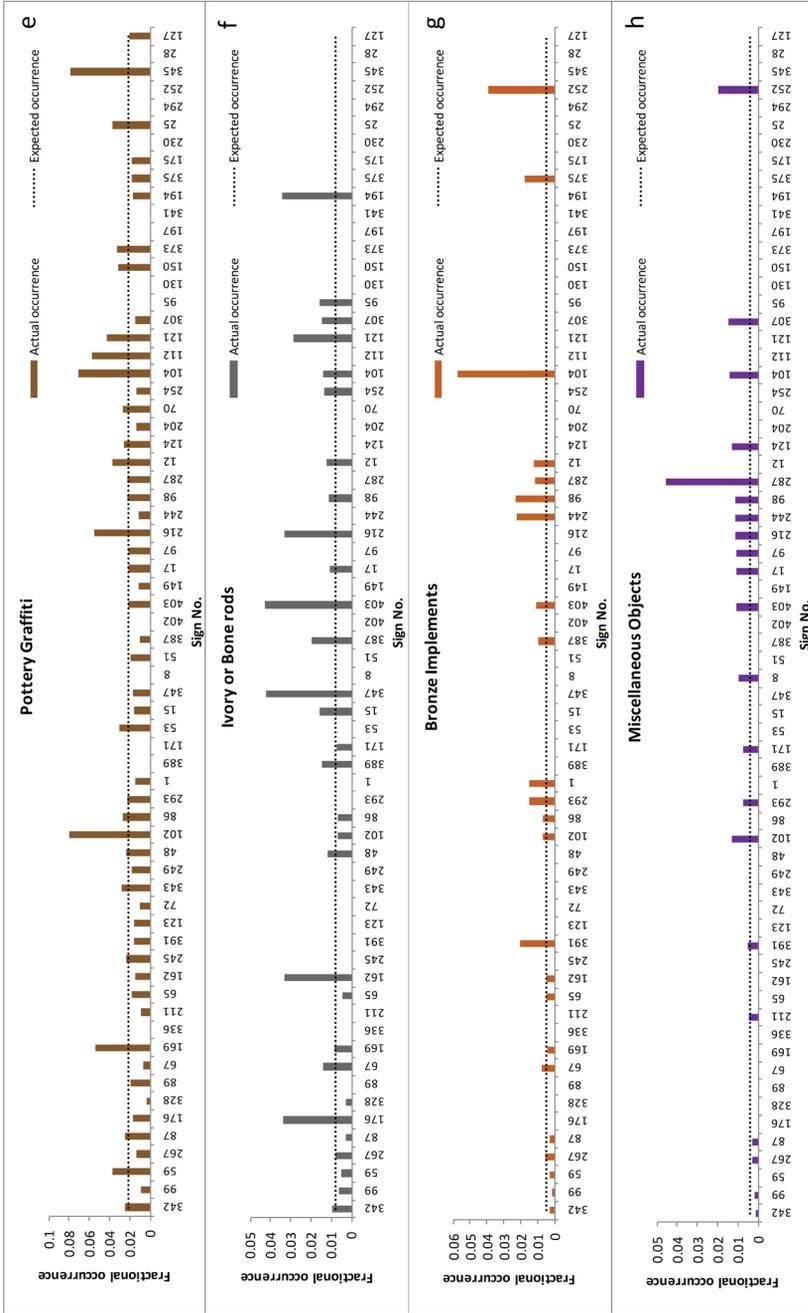


Figure 9. Fractional usage of the 67 most frequent signs on different types of objects (sorted in the descending order of their frequency in M77): (a) Seals, (b) Sealings, (c) Miniature tablets, (d) Copper tablets, (e) Pottery graffiti, (f) Ivory or bone rods, (g) Bronze implements, (h) Miscellaneous objects (See Table in the Appendix for images and frequency of signs marked on the x-axis). The line of expected occurrence for each type of object is derived from its sample size in M77.

frequent signs occur only on seals and sealings. Not all of these 67 most frequent signs occur on other types of objects. However, the percentage contribution of these signs to the total sample size of each type of object is close to 80% in all cases except for bronze implements and miscellaneous object types where they account for 61% and 52% of their total sample size respectively. The signs that are unique to various types of objects have low frequency of occurrence in M77 and their percentage contribution on different types of objects does not exceed 6%. For example, 169 signs out of 417 signs in the sign list of M77 occur only on seals, but as these signs have low frequency of occurrence, their total percentage contribution to the seals dataset is just 6%. The fractional usage of the 67 most frequent signs in M77 on different types of objects is given in Fig. 9.

As can be seen from Fig. 9 several signs deviate from their expected occurrences on different type of objects. The five signs on each type of object having highest deviation from their expected occurrences in the positive and negative directions are listed in Tables 6a and 6b respectively.

Following features are evident from Fig. 9 (a to h):

- 1) The sign usage on seals is fairly uniform except for sign numbers 95, 328 and 176 which deviate significantly from their expected occurrences and are amongst the dominant signs on miniature tablets (see section 5).
- 2) The usage of signs on sealings is close to their expected occurrence except for a few signs such as sign numbers 97, 124 and 194, which have significantly low frequency of occurrence.
- 3) The sign usage on copper tablets fluctuates dramatically around the expected line of occurrence. Most of the signs having high frequency in M77 appear with low frequency on copper tablets. Moreover, while the occurrence of sign numbers 169 and 389 is higher than their expected occurrence on copper tablets, sign numbers 162 and 387 are conspicuous by their absence. This is noteworthy as the sign numbers 169 and 389 share a high level of similarity in their design to the sign numbers 162 and 387 respectively (see Appendix).
- 4) Most frequent signs in pottery graffiti include sign numbers 102, 345 and 104.
- 5) Most frequent signs on ivory or bone rods are sign numbers 403, 347 and 194.
- 6) Sign numbers 104 and 252 form an overwhelming fraction of signs seen on bronze implements.

Table 6a. Signs having more than expected occurrence on various types of objects. The signs are listed such that the most frequent sign with highest excess is listed first.

S	SL	MT	CT	PG	IB	BI	Misc.
 28	 328	 95	 204	 102	 403	 104	 287
 123	 121	 328	 194	 345	 347	 252	 252
 17	 176	 176	 169	 104	 194	 98	 307
 391	 249	 48	 245	 112	 176	 244	 104
 70	 130	 89	 373	 216	 216	 391	 102

Table 6b. Signs having less than expected occurrence on various types of objects. The signs are listed such that the most frequent sign with highest deficit is listed first.

S	SL	MT	CT	PG	IB	BI	Misc.
 95	 97	 389	 176	 336	 89	 176	 59
 328	 124	 97	 162	 389	 336	 328	 176
 176	 194	 98	 391	 171	 211	 89	 328
 89	 70	 204	 343	 8	 245	 336	 89
 48	 12	 121	 86	 402	 391	 211	 67

4. Comparison of Data Between Sites and Object Types

In section 3, we have seen some empirical observations about the distribution of text lengths at various sites and on different types of objects. We also noticed preferential usage of signs at various sites and types of

objects. In this section, we compute the *distance*² (a measure of similarity or dissimilarity) between different sites and types of objects based on the following criteria

- 1) text length distribution, and,
- 2) usage of the 67 most frequent signs in M77.

4.1 Comparison of Sites

The similarity between various sites based on the text length distribution and usage of the 67 most frequent signs is plotted in Figs. 10a and 10b respectively.

As can be seen from Fig. 10a, Mohenjodaro and Lothal are closest with respect to their text length distributions and they form the cluster (MD, LL) which is quite distinct from other sites. The distance between Harappa and Kalibangan is slightly higher and they form another cluster (HP, KB). Chanhudaro clusters with Other Harappan sites and forms the cluster (CH, OH) though the distance between them is relatively higher compared to the sites in the two clusters (MD, LL) and (HP, KB). The clusters (HP, KB) and (CH, OH) are closer to each other than they are to (MD, LL) suggesting that the cluster (MD, LL) is very similar within itself but is quite different from other clusters. West Asian sites share least similarity with these clusters and it appears as a distinct entity in Fig. 10a.

In Fig. 10b, we compare the sites based on the usage of the most frequent 67 signs of M77. The pattern that emerges from the sign usage at various sites (Fig. 10b) is significantly different from the pattern emerging

² Distance (Y-axis) is a measure of the amount by which two entities differ and is computed using the method of hierarchical clustering. The links between the entities are represented as upside-down U-shaped lines. The height of the link indicates the distance between the entities. For example, in Fig. 10a, the link connecting the entities MD and LL has a height of 1.35 and the link connecting CH and OH has a height of 1.81. The height (or distance) of the link indicates the level of similarity between the entities connected by the link and it is inversely proportional to the level of similarity between entities. In Fig. 10a, the three clusters of sites, (MD, LL), (HP, KB) and (CH, OH), are joined by links of varying heights 1.35, 1.40 and 1.81. Even within these clusters, the level of similarity between entities of the cluster (MD, LL) is higher than the cluster (CH, OH) since the distance between the entities in the former cluster is less. Note that WA is quite distinct from rest of the sites, distance between them and WA being 3.71, the highest in this analysis.

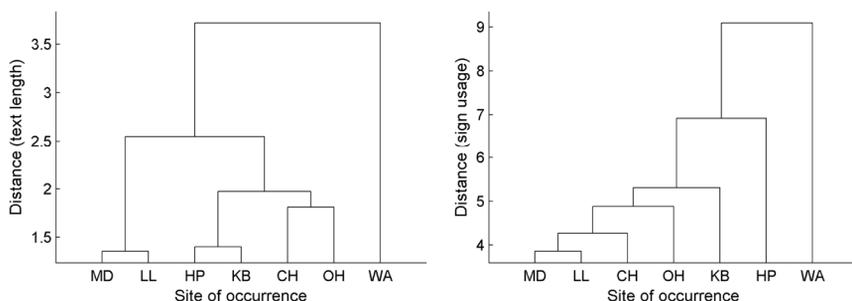


Figure 10. Similarity between sites based on a) text length, and, b) sign usage.

from the distribution of text lengths (Fig. 10a). Fig. 10b does not show the feature seen in Fig. 10a, namely some sites closest to some other sites form clusters and then they together get joined to another cluster at a higher distance.

It can be seen from Fig. 10b that Mohenjodaro and Lothal are also closest in terms of usage of signs and form a cluster (MD, LL). The site, which is closest to this cluster, is Chanhudaro followed by Other Harappan sites and Kalibangan in increasing order of distance. It is to be noted that the usage of signs at Harappa is significantly different not only from the cluster (MD, LL) but also from rest of the sites. Combined with Fig. 10a, this suggests that both in the text length distribution and in the usage of signs Harappa differs significantly from Mohenjodaro. We have seen in section 3.1 that the preferred type of object in Harappa is also significantly different from that of Mohenjodaro (Fig. 4). This suggests that Harappa has a distinct style of writing. However, we would like to emphasize that this does not in itself suggest that the content of writing in Harappa and Mohenjodaro are different as the same information can be expressed in different manner.

As in the case of text length distribution, the distance of West Asian sites is largest from all other sites in terms of sign usage. In our earlier work, we have shown that the grammar of the Indus texts on objects discovered from West Asian sites is also found to be significantly different from other Indus sites (Rao et al. 2009b) and the results here are consistent with our earlier observation.

Thus, it is clear from Figs. 10a and 10b that Mohenjodaro and Lothal share high level of similarity in their pattern of text length distributions

and usage of signs. While Harappa is closest to Kalibangan with respect to its text length distribution, this similarity does not hold in their pattern of sign usage. In terms of usage of signs, Harappa stands apart from all major sites.

4.2 Comparison of Object Types

As in the case of sites, we can compare various types of objects using the criteria of text lengths and sign usage. The similarity between various types of objects with regard to their text length distributions and the usage of the 67 most frequent signs is plotted in Figs. 11a and 11b respectively.

In Fig. 11a, we compare various types of objects based on their text length distributions. We find that, the distance between pottery graffiti and miscellaneous objects is least and they form the cluster (PG, Misc.). Sealings and miniature tablets form another cluster, however, the distance between them is higher than the objects in the cluster (PG, Misc.). Although the percentage of data from ivory or bone rods (IB) is very low, they are closest to seals in terms of their text length distribution forming the cluster (S, IB). The type of object that is closest to this cluster (S, IB) is copper tablets. Seals, ivory or bone rods and copper tablets together make a cluster that is quite different from other types of objects (Fig. 11a).

In Fig. 11b, we compare various types of objects based on the usage of the 67 most frequent signs of M77. We find that sealings and miniature tablets are closest to each other in terms of usage of signs forming the cluster (SL, MT). Moreover, seals share a high level of similarity with the pottery graffiti. Bronze implements are closest to miscellaneous objects in

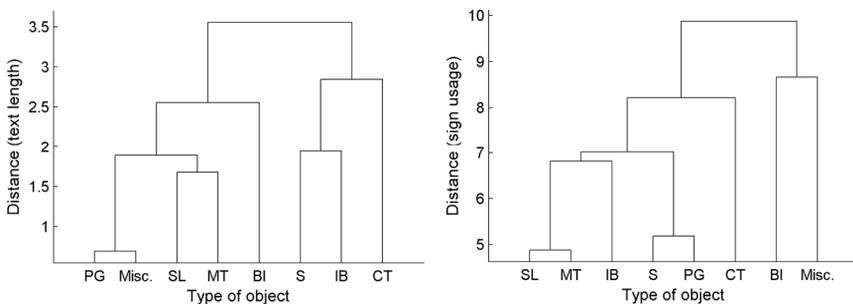


Figure 11. Similarity between objects based on a) text length, and, b) sign usage.

terms of their usage of signs and they together form a cluster that is quite dissimilar from other types of objects.

Thus, it is clear that the clustering of different types of objects is significantly different with respect to text lengths (Figs. 11a) and usage of signs (Fig. 11b) except for sealings and miniature tablets that form a cluster in both the cases. The similarity between seals and pottery graffiti is noteworthy. The present analysis suggests that seals and pottery graffiti are very similar in terms of the usage of signs, however, they are quite distinct in their text length distributions. A detailed comparison of their text length distributions (Fig. 7) suggests that seals are a more elaborate version of writing in pottery graffiti. Kenoyer (2006) has studied the evolution of various signs of the script and has suggested their origin as pottery graffiti. The results here seem to be consistent with this suggestion.

5. Frequent Texts: A Case Study

Four texts in M77 appear unusually large number of times and together account for about 8% of the total number of lines of texts and 4% of the total sign occurrences. These texts are (89, 328), (87, 328), (95, 328) and (48, 342, 176). The positional distribution of the sign combinations in these texts is given in Table 7.

It can be seen from Table 7 that the sign combinations appear mostly as solo (independent) texts and very rarely as part of longer texts. The

Table 7. Positional distribution of sign combinations in the four most frequent texts.

Serial No.	Sign Combination	Image	Frequency of occurrence				
			Solus	Initial	Medial	Final	Total
1	89, 328		111	3	0	10	124
2	87, 328		74	1	0	3	78
3	95, 328		50	4	3	1	58
4	48, 342, 176		41	1	0	4	46

Table 8. Distribution of the four most frequent texts with sites.

Serial No.	Text	Text image	Frequency of occurrence							Total
			MD	HP	LL	KB	CH	OH	WA	
1	89, 328		0	111	0	0	0	0	0	111
2	87, 328		0	73	0	1	0	0	0	74
3	95, 328		0	50	0	0	0	0	0	50
4	48, 342, 176		1	39	0	1	0	0	0	41

Table 9. Distribution of the four most frequent texts with object types.

Serial No.	Text	Text image	Frequency of occurrence							Total	
			S	SL	MT	CT	PG	IB	BI		Misc.
1	89, 328		0	49	62	0	0	0	0	0	111
2	87, 328		0	37	37	0	0	0	0	0	74
3	95, 328		0	9	41	0	0	0	0	0	50
4	48, 342, 176		0	3	37	0	0	1	0	0	41

distributions of these four texts with sites and types of objects are given in Tables 8 and 9 respectively.

As can be seen from Tables 8 and 9, almost all occurrences of these texts are on sealings and miniature tablets from Harappa. At Harappa, these texts account for about 23% of the texts on sealings and 36% of the texts on miniature tablets. These are usually two or three sided objects (prism-shaped), sometimes with texts written on both sides and a field symbol (often one or more circles with a dot) on the third side. In order to explore the relation between these frequent texts we analysed all the texts that appear on the other side of the objects having the text (48, 342, 176) on one side. In Fig. 12 we provide a Venn diagram of the four most frequent texts.

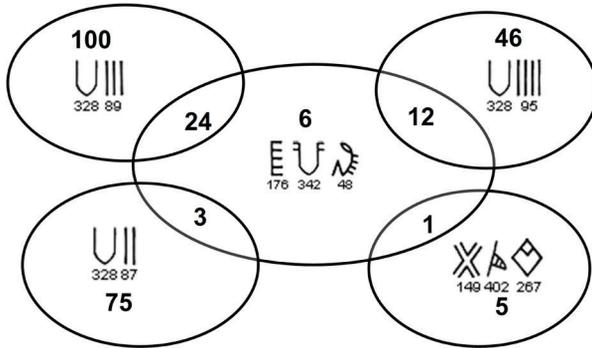


Figure 12. Venn diagram of the four most frequent texts in M77. The number of times a pair of texts appears on different sides of objects is marked in the overlapping region of the Venn diagram.

It can be seen from Fig. 12 that the text (48, 342, 176) seems to be a common thread for the texts (89, 328), (95, 328), (87, 328) and (267, 402, 149) occurring on other side of these objects. For instance, the text (89, 328) appears 24 times with the central text (48, 342, 176) and 100 times with others (Fig. 12). Furthermore, the texts (89, 328), (95, 328), (87, 328) do not appear back to back on the same object.

Thus objects with the text (48, 342, 176) on one side have either of the texts (89, 328), (95, 328) and (87, 328) on the other side in different instances.³ The latter texts share similarity in their structure and seem to have some contextual relation with the text (48, 342, 176).

6. Discussion and Conclusions

The Indus valley civilisation was a complex civilisation with such a humungous spread that it could not have been homogenous in space or time (Vahia and Yadav 2011). One of its most mysterious creations is its script that has defied decipherment in spite of more than a century of efforts. This is mainly due to the paucity of data, absence of bilingual

³ There are two cases when the miniature tablets containing the text (48, 342, 176) have writing on two other sides. In one such case the two other texts are (89, 328) and (89, 328). In the other case, the two other texts are (267, 402, 149) and (89, 328).

or multilingual texts, lack of knowledge of language(s) spoken by the Indus people and apparent discontinuity of traditions at the decline of the civilisation.

Various types of inscribed objects containing the Indus script have been discovered at several sites of the Indus valley civilisation that are often separated by hundreds of kilometres. Around 85% of these objects have been excavated from Mohenjodaro and Harappa alone, the two major sites of the Indus valley civilisation and remaining 15% of the objects are from other sites (Table 1). Moreover, about 80% of these objects are in the form of seals and sealings and remaining types of objects (such as miniature tablets, copper tablets, pottery graffiti, ivory or bone rods, bronze implements etc.) comprise about 20% of the data in M77 (Table 2). A comprehensive study of the Indus script therefore needs a far more holistic approach so that it can be related to its contextual parameters and its content or implied meaning can be understood (Fig. 2).

One of the outstanding features of the Indus script is its compact format (the inscribed objects are generally a few square centimetres in size) and its large spread. We know that the Indus valley civilisation was extensive in its reach and it sourced raw materials from vast distances (Law 2011). Moreover, samples of objects inscribed with the Indus script have been found at most of its sites. The spread of the script over large distances suggests that the level of coding and standardisation of the script must have been very precise and most of the script, within a given context (such as a specific type of object) may have permitted only small fluctuations. This is well borne out of our study.

The Indus writing seems to be a complex mix of multiple pieces of information, of different times and of different kinds of records. Our study of the structural aspects of the script (Yadav 2012, Yadav et al. 2010, Rao et al. 2010, Rao et al. 2009a, b, Yadav et al. 2008a, b) suggests that the Indus texts have an underlying logic in their structure and seem to be dictated by certain set of rules. Even a first order (bigram) model of the Indus script is able to predict missing signs in the texts with about 75% accuracy (Yadav et al. 2010, Rao et al. 2009b). However, the significant difference in the rules of the written material found at West Asian sites from that of the Indus sites suggests that the script was versatile enough to code different content (Rao et al. 2009b).

Yadav and Vahia (2011b) have shown that about 63% of the signs in

the sign list are 'composite signs' that is, they are created by either merger of two or more 'basic' signs or by adding special 'modifiers' to the 'basic' signs. However, the high frequency of usage of some of the composite signs and their evolution into more complex variants suggests that, once identified, these composite signs became part of the core signs. The usage of the 'modifiers' at almost all sites also suggests that the manner of modifying the signs was universally agreed upon over the entire area of the Indus valley civilisation (Yadav and Vahia 2011b).

We therefore need to understand the level of non-uniformity in the Indus script against conditions where uniformity seems to be the norm. For instance, although the set of the 67 most frequent signs which accounts for 80% of the data in M77, often accounts for similar percentage of data for all sites and types of objects (Tables 3 and 5), the relative contribution of each of these 67 signs fluctuates across various sites and types of objects (Figs. 8 and 9). It is therefore important to study the fluctuations in various aspects of the Indus script (such as text lengths or sign usage) with respect to different contextual parameters (such as sites or object types).

Our study clearly demonstrates that the type of object, the length of the written texts and the usage of signs are a complex set of parameters that are sensitive to the sites from which the writing is discovered. There are small and subtle variations across sites and types of objects but no large and dramatic variations suggesting that the level of standardisation was large. The small variations suggest that even within this framework, some local freedom was permitted. This shows that the Indus valley civilisation, though highly standardised in several aspects (as seen from other evidence, see for e.g. Vahia and Yadav 2011) was not fully homogenised.

In particular, we find that the largest site, Mohenjodaro seems to have a bias towards the seals while Harappa seems to have a bias towards sealings and miniature tablets (Fig. 4). It is normally assumed that sealings (being impressions of seals) were created using seals and hence we must find the seal for every sealing. However, while it has been known that not many seals corresponding to the sealings have been found, the statistical study here shows that in fact, in terms of text length distribution and usage of signs, seals and sealings are significantly different (Figs. 11a, 11b). This implies that the seals used to make these sealings are not present in the set of objects that have survived and discovered from various sites.

The apparent preference of texts of length five and eight in some of

the cases is also significant as is the fact that the texts of lengths 9 to 14 are found exclusively on seals and sealings. Yadav et al. (2008b) have segmented longer texts by identifying text enders, text beginners and frequent sign combinations. They have shown that most of the longer texts seem to be a combination of smaller texts, something that was anecdotally suggested by Mahadevan (1978) and Parpola (1994). Moreover, there are important site and object type specific differences in sign usage, which cannot be, overlooked (Tables 4a, 4b and 6a, 6b). Any model of interpretation of the Indus script must therefore include site of occurrence and type of object as potential sources of fluctuations.

The differences in the results derived from the comparison of text length distributions and the usage of signs across different sites and types of objects should be studied carefully. The dissimilarity between Mohenjodaro and Harappa and the similarity between Mohenjodaro and Lothal remains unaffected with respect to both these parameters (Figs. 10a and 10b). Although Harappa is closest to Kalibangan with respect to its text length distribution, it is quite distinct in its sign usage not only from Kalibangan but also from all other sites (Figs. 10a and 10b). The uniqueness of the data from West Asian sites is also evident from this study.

In case of objects, the distance between sealings and miniature tablets is least in terms of their usage of signs. Seals are found to be closest to pottery graffiti with respect to usage of signs and this result is consistent with Kenoyer (2006). The study suggests the origin of the signs of the Indus script in pottery graffiti and shows that beginning with a much larger set of strokes in pottery graffiti, the Indus script in its final format uses some select number of signs. The question addressed here is whether there were further preferences in the use of signs on different types of objects and at different sites.

It is also possible that some of the signs in the sign list of M77 could be stylistic variants of the same sign. This is implied by the high frequency of occurrence of the signs, sign numbers 169  and 389  on copper tablets and the complete absence of visually similar signs, sign numbers 162  and 387  on copper tablets. A detailed analysis of the stylistic variations of the signs across object types and sites can throw more light on this. With the present framework, it should now be possible to check such suggestions more vigorously by doing site and object type sensitive analysis with a prior knowledge of affinity brought out by the study.

The Indus valley people had attained a high level of uniformity in various aspects of their lives. Moreover, the Indus writing was not an act of casual creation but involved a complex activity with a carefully designed script (Yadav and Vahia 2011b) and often, extensive addition of non-script material (field symbols) on the inscribed objects. While there is a common thread of rules and grammatical structures that are fairly well obeyed in the Indus writing, the present study suggests that writing on different types of objects and at different sites do have individualistic clues to their content. Further studies on the use of various motifs on the inscribed objects, the shifts in the designs of signs and stratigraphy will add more clarity to this extremely complex problem.

Appendix

67 most frequent signs with their frequency in M77.
(Sorted in decreasing order of frequency)

Sign	Freq	Sign	Freq	Sign	Freq	Sign	Freq	Sign	Freq
 342	1395	 391	195	 8	105	 70	73	 230	54
 99	649	 123	193	 51	105	 254	73	 25	53
 59	381	 72	188	 387	102	 104	70	 294	53
 267	376	 343	177	 402	99	 112	70	 252	51
 87	365	 249	170	 403	93	 121	70	 345	51
 176	355	 48	168	 149	92	 307	69	 28	50
 328	323	 102	151	 17	91	 95	64	 127	50
 89	314	 86	149	 97	91	 130	63		
 67	279	 293	136	 216	90	 150	63		
 189	240	 1	134	 244	89	 373	61		
 336	236	 389	134	 98	88	 197	60		
 211	227	 171	132	 287	88	 341	59		
 65	216	 53	130	 12	80	 194	58		
 162	212	 15	126	 124	78	 375	57		
 245	207	 347	118	 204	76	 175	54		

References

- Agrawal, D. P. (2007). *The Indus Civilization: An Interdisciplinary Perspective*. New Delhi: Aryan Books International.
- Agrawal, D. P. (2009). *Harappan Technology and its Legacy*. Delhi: Rupa and Co. and Infinity Foundation Series.
- Joshi J. P. and Parpola A. (1987). *Corpus of Indus Seals and Inscriptions, 1. Collections in India*, Suomalainen Tiedekatemia, Helsinki, also Memoirs of the Archaeological Survey of India no. 86.
- Kenoyer, J. M. (1998). *Ancient Cities of the Indus Valley Civilization*. Oxford: Oxford University Press.
- Kenoyer, J. M. (2006). The Origin, Context and Function of the Indus Script: Recent Insights from Harappa. *Proceedings of the Pre-symposium and the 7th ESCA Harvard-Kyoto Roundtable*. T. Osada and N. Hase. Kyoto, Research Institute for Humanity and Nature, RIHN. 9–27.
- Koskenniemi, K. And Parpola A. (1982). *A Concordance to the Texts in the Indus Script*, Department of Asian and African Studies, University of Helsinki, Research Reports, No. 3.
- Law, R. W. (2011). Inter-Regional Interaction and Urbanism in the Ancient Indus Valley: A Geologic Provenience Study of Harappa's Rock and Mineral Assemblage. *Occasional Paper 11, Linguistics, Archaeology and the Human Past*. Indus Project, Research Institute for Humanity and Nature, Kyoto.
- Mahadevan, I. (1977). *The Indus Script: Texts, Concordance and Tables*, Memoirs of the Archaeological Survey of India No. 77.
- Mahadevan, I. (1978). Recent Advances in the Study of the Indus Script, *Puratattva* 9. 34–42.
- Mahadevan, I. (2002). Aryan or Dravidian or Neither? A Study of Recent Attempts to Decipher the Indus Script (1995–2000). *Electronic Journal of Vedic Studies (EJVS)* 2. 1–23.
- Parpola, A. (1994). *Deciphering the Indus Script*. Cambridge: Cambridge University Press.
- Parpola, A. (2005). Study of the Indus Script. *Transactions of the International Conference of Eastern Studies* 50. 28-66. Tokyo: The Tôhō Gakkai.
- Parpola A., Pande B. M., and Koskikallio Petteri (2010). *Corpus of Indus*

- Seals and Inscriptions, Volume 3: New Material, Untraced Objects, and Collections Outside India and Pakistan*, Suomalainen Tiedeakatemia, Helsinki.
- Possehl, G. L. (1996). *Indus Age: The Writing System*, Oxford IBH Publishing Co. Pvt. Ltd.
- Possehl, G. L. (1999). *Indus Age: The Beginnings*. Philadelphia: University of Pennsylvania Press; New Delhi: Oxford & IBH Publishing Co. Pvt. Ltd.
- Possehl, G. L. (2002). *The Indus Civilization: A Contemporary Perspective*, Vistaar Publications, New Delhi.
- Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., and Mahadevan, I. (2009a). Entropic Evidence for Linguistic Structure in the Indus Script, *Science* **324**. 1165.
- Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., and Mahadevan, I. (2009b). A Markov Model of the Indus Script, *Proceedings of the National Academy of Sciences* **106**. 13685-13690.
- Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., and Mahadevan, I. (2010). Entropy, the Indus Script and Language: A Reply to R. Sproat, *Computational Linguistics* **36(4)**. 795-805.
- Shah, S. G. M. and Parpola, A. (1991). *Corpus of Indus Seals and Inscriptions, 2 Collections in Pakistan*, Suomalainen Tiedeakatemia, Helsinki, also *Memoirs of the Archaeology and Museums, Government of Pakistan* Volume 5.
- Sinha, S., Yadav, N., and Vahia, M. N. (2011). In Square Circle: Geometric Knowledge of the Indus Civilization, *Math Unlimited: Essays in Mathematics*, eds. R. Sujatha, H. N. Ramaswamy and C. S. Yogananda, Science Publishers, USA. 451–462.
- Vahia, M. N., and Yadav, N. (2010). Harappan Geometry and Symmetry: A Study of Geometrical Patterns on Indus Objects, *Indian Journal of History of Science* **45.3**. 343–36.
- Vahia, M. N. and Yadav, N. (2011). Reconstructing the History of Harappan Civilisation, *Journal of Social Evolution and History* **10(2)**. 67–86.
- Wells, B. K. (2006). *Epigraphic Approaches to Indus Writing*, Ph.D. Thesis, Harvard University, Cambridge, Massachusetts.
- Wright, R. P. (2010). *The Ancient Indus – Urbanism, Economy and Society*. New York: Cambridge University Press.
- Yadav, N. (2012). Statistical Studies of the Indus Script, *Man and Environment XXXVII(1)*. 1–7.

- Yadav, N. and Vahia, M. N. (2011a). Classification of Patterns on Indus Objects, *International Journal of Dravidian Linguistics* **40**. 89–114.
- Yadav, N. and Vahia, M. N. (2011b). Indus Script: A Study of its Sign Design, *Scripta* **3**. 133-172.
- Yadav, N., Joglekar, H., Rao R. P. N, Vahia, M. N., Adhikari, R., and Mahadevan, I. (2010). Statistical Analysis of the Indus Script using n-grams, *PLoS One*, Vol. 5, Issue 3. <http://dx.plos.org/10.1371/journal.pone.0009506>.
- Yadav, N., Vahia, M. N., Mahadevan, I., and Joglekar, H. (2008a). A Statistical Approach for Pattern Search in Indus Writing, *International Journal of Dravidian Linguistics* **XXXVII(1)**. 39–52.
- Yadav, N., Vahia, M. N., Mahadevan, I., and Joglekar, H. (2008b). Segmentation of Indus Texts, *International Journal of Dravidian Linguistics* **XXXVII(1)**. 53–72.

Nisha Yadav
Tata Institute of Fundamental Research
Mumbai-400 005, India
<y_nisha@tifr.res.in>

[Received 29 October 2012;
revision received 6 August 2013;
accepted 28 August 2013]

