Indus Script: A Study of its Sign Design

Nisha Yadav* and M. N. Vahia
Tata Institute of Fundamental Research

The Indus script is an undeciphered script of the ancient world. In spite of numerous attempts over several decades, the script has defied universally acceptable decipherment. In a recent series of papers (Yadav et al. 2010; Rao et al. 2009a, b; Yadav et al. 2008a, b) we have analysed the sequences of Indus signs which demonstrate presence of a rich syntax and logic in its structure. Here we focus on the structural design *of individual signs of the Indus script. Our study is based on the sign list given in the concordance of Mahadevan (1977) which consists of 417 distinct signs. We analyse the structure of all signs in the sign list of Indus script and visually identify three types of design elements of Indus signs namely basic signs, provisional basic signs and modifiers. These elements combine in a variety of ways to generate the entire set of Indus signs. By comparing the environment of compound signs with all possible sequences of constituent basic signs, we show that sign compounding (ligaturing) and sign modification seem to change the meaning or add value to basic signs rather than save writing space. The study aims to provide an understanding of the general makeup and mechanics of design of Indus signs.*

Keywords: *Indus script, Harappan script, ancient scripts, undeciphered scripts, sign design, structure of Indus signs, sign compounding, ligatures*

1. Introduction

Writing is an important window to the intellectual creativity of a civilisation. Renfrew points out that the practice of writing and the development of a coherent system of signs — a script — is something that is seen only in complex societies and calls it a feature of civilisations (Renfrew 1989: 20). Houston (2004) addresses several issues related to the

* Address for correspondence: y_nisha@tifr.res.in

SCRIPTA, Volume 3 (June 2011): 1-36

script development and origins of writing in cultural context. The history of writing and the evolution of several alphabetic and non-alphabetic scripts are also discussed in Daniels and Bright (1996) and references therein. Damerow (2006) has examined issues related to monogenesis and polygenesis of writing and presents arguments on both sides without committing to a specific point of view. However, considering the variations in different writing systems, writing probably was invented independently at several places.

Changizi et al. (2006) have analysed the design configuration of signs across various types of visual signs including a set of 96 non-logographic writing systems, Chinese as well as non-linguistic systems and they conclude that the design of signs is greatly influenced by the environment. This suggests that even if the idea of writing is acquired, the script and the configuration of signs reflect strong local influences. When a script is more complex than pictographs, the association between the design of a sign, meaning of the sign and its abstraction becomes highly involved and implies specific meaning that is accepted by all those who use that script. Even the simplest, pictographic scripts require a certain agreement between the users on the association between the pictograph and the associated object.

Decipherment of a script on the other hand is a different problem altogether. It is often aided by the discovery of a multilingual text where the same text is written in an undeciphered script as well as known script(s). Both Egyptian hieroglyphs and Mesopotamian cuneiform script were deciphered with the help of multilingual finds. In some cases, continuing linguistic traditions provide significant clues and at times interlocking phonetic values are used as a proof of decipherment (Pope 1999). In the absence of these, statistical studies can provide important insights into the structure of the writing and can be used to define a syntactic framework for the script (see Yadav et al. 2010).

The Indus script is a product of one of the largest bronze age cultures (Wright 2010, Agrawal 2007, Possehl 2002, Kenoyer 1998). At its peak from 2500 BCE to 1900 BCE, the civilisation was spread over an area of about a million square kilometres across most of the present day Pakistan and north-western India. It was known to have about 1500 settlements (Kenoyer 1998: 17) with several large urban centres. It was distinguished for its highly utilitarian and standardised life style, excellent water

management system and architecture. The Indus script is predominantly found on objects such as seals and sealings, and it also makes its appearance on other objects such as copper tablets, ivory sticks, bronze implements, pottery etc. from almost all sites of this civilization. A comprehensive photographic documentation of Indus seals and inscriptions is available as three volumes of *Corpus of Indus Seals and Inscriptions* (henceforth referred to as CISI 1-3). The Indus script has defied decipherment in spite of several serious attempts (see Possehl 1996 for a critical review of various attempts). Robinson (2009) provides a more recent perspective on several undeciphered scripts of the world. He also discusses the issues and complexities related to the problem of decipherment of Indus script and refers to it as the biggest challenge in archaeological decipherment (Robinson 2009: 265). This is primarily because no multilingual texts have been found and the underlying language(s) is unknown. In addition to these, the problem becomes more challenging because the script occurs in very short texts. The average length of an Indus text is five signs and the longest text in a single line has only 14 signs (Mahadevan, 1977).

The nature and content of the Indus script has been extensively debated in the literature. More than a hundred attempts have been made to assign meanings to various signs and sign combinations, relating it to proto-Dravidian language (see Parpola 2009, 1994, Mahadevan 1998) on the one hand and to Sanskrit (Rao 1982) on the other. It has even been suggested that the script is entirely numeric (Subbarayappa 1997). However, no consistent and generally agreed interpretation exists and most interpretations are at variance with each other and, at times, internally inconsistent (Possehl 1996). Parpola (2005) and Mahadevan (2002) provide a more recent perspective on the study of Indus script. The statistical approach has been explored by us and others elsewhere (see Yadav et al. 2010; Rao et al. 2009a, b; Yadav et al. 2008a, b; Parpola 1994, Siromoney and Haq, 1988) to investigate the structure of the Indus script. A series of mathematical tests on the manner of sequencing of Indus signs (Yadav et al. 2010, Rao et al. 2009a, b, Yadav et al. 2008a, b) makes it clear that the Indus writing is highly ordered and has a specific grammar to it. Though no independent evidence is yet available to check if the writing is linguistic or not, some interesting similarity in the flexibility of its sign usage with linguistic systems have been reported (Rao et al. 2009b, 2010). However, the *design* of the Indus signs has received little attention except for a few

studies that are summarised below.

Kenoyer (2006) has studied the evolution of signs from early to mature Harappan period by a careful study of the stratigraphy and quality of engraved material excavated from Harappa. The study has shown that the writing becomes more standardised and uniform as the culture evolves into its mature period (2600–1900 BCE). It also shows that a fraction of the pottery markings found in early period seem to have evolved into signs in the sign list of Indus script. The study shows that the script goes from simple pottery markings during the Ravi phase (3900–2800 BCE) through Kot Diji phase (3900–2600 BCE) and becomes standardised in the mature Harappan phase (2600–1900 BCE).

The study of the design of Indus signs and their modifiers has received attention only in the context of understanding the effect of some specific modifiers (Mahadevan 2006) or for some select examples (Parpola 1994, Wells 2006). The classification of signs into basic and composite has been suggested by Parpola (1994) where he explores the effect of ligaturing using some select examples of Indus texts. By comparing the occurrence pattern of some of the composite signs with the occurrence pattern of their basic version, he suggests that that some of the components in composite signs seem to be phonetic or semantic determinatives (Parpola 1994: 79). More recently, Wells (2006, Table 3.4 p 80) has also attempted to classify the signs into several categories such as simple signs, signs with elaboration, compound signs, enclosures, strokes, multiple class etc. Again, by comparing the contexts of a few compound signs with the constituent simple signs Wells (2006: 83) suggests that compounding seem to signal a shift in semantic values of the constituents.

In the present study, we undertake a comprehensive analysis of the design of Indus signs. Here we attempt to analyse the structural design of the Indus signs and the interrelation between their components. The entire analysis is independent of any linguistic or interpretative model but any attempt to understand the writing *must* address the question regarding how the signs are designed and modified. We first identify the various design elements namely basic signs, provisional basic signs and modifiers and, using these elements, decompose all the composite signs. We then quantify the differences in the occurrence pattern of compound signs (a class of composite signs) with all possible sequences of their original constituents by comparing their environment in terms of the signs preceding or

Figure 1. A large unicorn seal from Harappa on the left (Copyright Harappa Archaeological Research Project/J.M. Kenoyer, Courtesy Dept. of Archaeology and Museums, Govt. of Pakistan) and some Indus signs from the sign list of M77 on the right. The seal is about 5 cm x 5 cm in size.

following them.

We use the sign list of Mahadevan (1977; henceforth referred to as M77) for analysis. In Fig. 1 we have given an image of an Indus seal from Harappa and a sample of Indus signs from the sign list of M77. There has been about 10% increase in the size of the corpus as a result of recent excavations since M77 was created and a few new signs have emerged (see CISI 3). There has also been a re-evaluation of the issues such as the identification of a sign as a distinct sign or as a variant. For example, some of the sign variants noted in M77 are considered as different signs in Wells (2006). However, Parpola (1994) suggests about 400 signs in the sign list and is in general accordance with the signs listed in M77 with a few exceptions. These variations over the original construct of signs listed in M77 are unlikely to alter the basic conclusions here.

2. Dataset

The total number of signs in the Indus script is generally agreed to be around 400 (Parpola 1994, Mahadevan 1977), though Wells (2006)

identifies about 676 distinct signs. In the present work, we investigate the design of the 417 distinct signs identified 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 the present work aims to get a broad understanding of the design of Indus signs, the issue of variants and total number of signs are not taken into account. While the subtle differences in the writing of individual signs and their regional variations which may carry additional information, are lost in the normalised sign lists of all concordances, the normalised sign list can still provide a broad understanding of the general makeup of the signs. The present study also does not take into account variations in sign design due to site of occurrence, stratigraphy, type of objects and quality of writing. With these caveats, we have used the normalised sign list of M77 to understand the general structure of Indus signs.

M77 has organised the signs in the sign list based on their visual similarity and sequenced them from sign number 1 to 417 in that order. The serial number of the signs used in this paper is as given in M77. In listing and numbering the various signs of Indus script, M77 follows a certain order in clustering similar looking signs which, in a broad sense, goes from simple to complex design in each category. The numbering of the signs therefore has only approximate correlation with similar looking signs but the association is arbitrary. Hence proximity of sign numbers does not necessarily indicate similarity in design or function and each number label is taken as an independent entity. 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 by just strings of sign numbers are to be read from left to right.

2.1 Criteria for sign classification and decomposition

In order to classify signs based on their design we analyse each sign and *visually identify* various elements that seem to have been included in its design. In principle, the design of all signs consists of strokes and they can all be decomposed into vertical and horizontal strokes and curves with a few inclined curves or lines. However, such decomposition would lead to a chaotic mess where the idea of what is a unit of information will be lost. Conversely, keeping the signs as they appear in the sign list will not allow

any insights into the manner in which these sign components have been used. We have therefore adopted a set of criteria which allows us to judge whether a sign is a basic sign or a composite sign. The criteria we follow are:

1) A sign is considered as a basic sign if it has a simple geometric design and cannot be decomposed further into identifiable units. Signs such as sign numbers 1, 328, 373, 176 (Fig. 2a) are included in this category. For example, we do not decompose sign number 1 into a vertical stroke and four inclined strokes since the sign is a basis for several other signs. Many signs of complex design, which have very rare occurrence (one or two) and are not obviously composites of basic signs are also considered as basic signs.
2) The signs that are identified as basic signs have one or more of the following characteristics:
   a. It has a high frequency of occurrence.
   b. It has other basic signs merging into it.
   c. It appears with one or more modifiers.
   d. Its environment is different than the sign from which it seems to be derived.

One example of this is sign number 342 which appears to be a derivative of sign number of 328 in terms of its design. However, sign number 342 is the most common text ender and is also the sign with the highest frequency of occurrence in M77. Hence, we assume that even if the design of sign number 342 may have been influenced by the design of sign number 328 , they seem to have different functions. For these reasons, 342 is retained as a basic sign. This classification of 342 as a basic sign is further reinforced by the fact that it is also used in the design of several composite signs such as sign numbers 15 , 352 , 353 and 394 .

The decomposition of signs is done with a consistent logical assumption which was decided before hand and applied uniformly across all the signs in the sign list. The criteria assigns no value, meaning or interpretation to a sign but gauges the status, environment and usage of a sign to evaluate if the sign is of core importance to the writers and if it is ligatured with other basic signs for whatever reasons that they may have had.

3. Types of Indus signs based on their design

A careful look at the design of each sign can help us classify the sign into one of the two broad categories. They are basic signs and composite signs. We discuss each of these below.

3.1 Basic signs

Basic signs are signs that cannot be decomposed further into simpler signs and/or form elements of composite signs. A total of 154 signs in the sign list of M77 are identified as basic signs. Most frequently used basic signs in the design of Indus signs are sign numbers 1, 328, 373 and 176 (Fig. 2a).

3.2 Composite signs

Composite signs are signs that can be decomposed into one or more of the three design elements defined below. A total of 263 signs in the sign list of M77 are identified as composite signs. Some examples of composite signs are given in Fig. 2b.

3.2.1 Design elements of composite signs

The three design elements of composite signs are: basic signs, provisional basic signs and modifiers. They are defined as below:

Figure 2a. Examples of basic signs.

Figure 2b. Examples of composite signs.

1) Basic signs

As stated earlier (in section 3.1, Fig. 2a), basic signs are signs that cannot be decomposed further into simpler signs and/or form elements of composite signs.

2) Provisional basic signs

In the design of several composite signs we can identify elements that are ligatured to other basic signs or modifiers. However, these elements do not appear as distinct signs in the sign list of M77. We refer to these elements as provisional basic signs. Provisional basic signs are similar to basic signs in terms of their characteristic usage in the design of composite signs. However, unlike basic signs they do not have independent existence. The fact that they have not been seen as independent signs may be due to sample incompleteness and it is likely that at some later date texts with these signs are found. It is also possible that they may not have been very useful as distinct signs and were not used independently. We identify 10 provisional basic signs that are referred to by numbers 801 to 810. Table 1 lists all provisional basic signs along with all the signs in which they appear and their total frequency of occurrence in different signs.

3) Modifiers

Modifiers are the design elements that do not have independent existence but appear at specific locations in composite signs. They modify the basic signs in a variety of ways to create the composite signs. We identify 21 modifiers such as '^', '| |' etc. that are added to the basic signs at various locations and are referred to by numbers 901 to 921 Fig. 3. Table 2 lists all modifiers along with all the signs in which they appear and the total frequency of occurrence of each modifier on different signs.

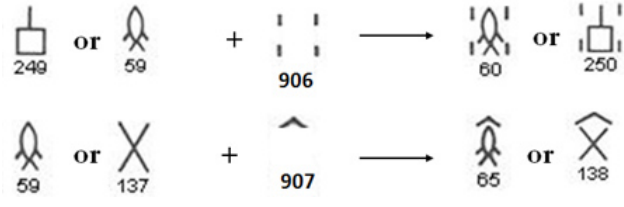
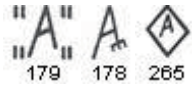
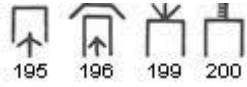
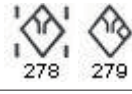
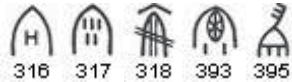


Figure 3. Modification of basic signs using modifiers 906 and 907.

Table 1. List of provisional basic signs along with their occurrences.

Serial No.	Provisional basic sign	Description	Design	Sign(s) in which it appears	Total Freq.
1	801	Wobbly legs		 44	1
2	802	Cross legged		 45 46	2
3	803*	'A'		 179 178 265	3
4	804	Inverted box		 195 196 199 200	4
5	805	Two legged Y		 278 279	2
6	806	"VA" joined		 131	1
7	807	Inverted U		 316 317 318 393 395	5
8	808	H		 316	1
9	809	Shaded fish-like shape without fins		 75	1
10	810	Vertical line with a bulb at bottom		 34 335 336 337 354	5

* The provisional basic sign 803  is listed as a variant of sign number 178 in M77.

Table 2. List of modifiers along with their occurrences (Note that at times one or more modifiers may appear more than once in the same sign).

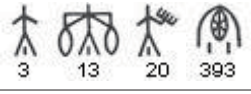
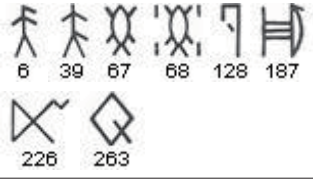
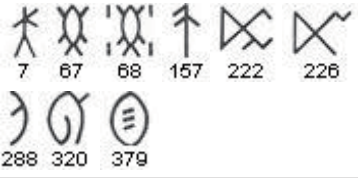
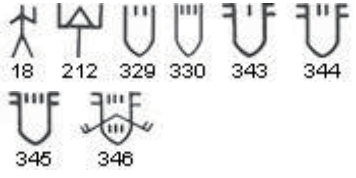
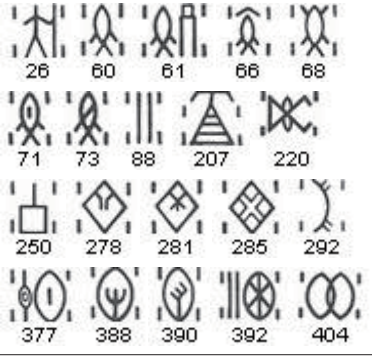
Modifier	Description	Design	Sign(s) in which it appears	Total Freq.
901	Vertical stroke on both sides		 2 291 401 416	4
902	Vertical stroke at the bottom		 3 13 20 393	5
903	Right downward tilted stroke		 6 39 67 68 128 187 226 263	8
904	Right upward tilted stroke		 7 67 68 157 222 226 288 320 379	11
905	Vertical stroke at the top		 18 212 329 330 343 344 345 346	20
906	Four vertical strokes enclosing the sign		 26 60 61 66 68 71 73 88 207 220 250 278 281 285 292 377 388 390 392 404	20

Table 2. (continued)

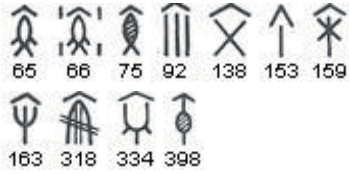
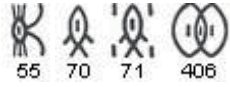
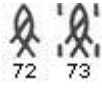
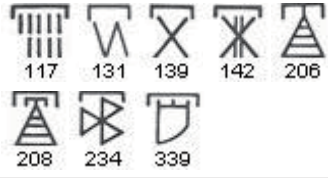
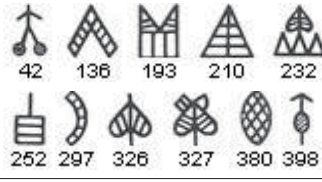
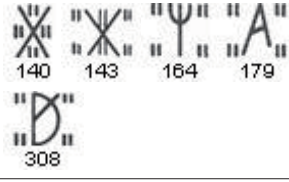
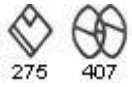
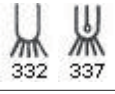
Modifier	Description	Design	Sign(s) in which it appears	Total Freq.
907	Pointed Hat		 65 66 75 92 138 153 159 163 318 334 398	11
908	Small central vertical stroke in the centre		 55 70 71 406	6
909	Slanted line on both sides		 85	1
910	Slanted line in the centre		 72 73	2
911	Flat hat		 117 131 139 142 206 208 234 339	9
912	Shading		 42 136 193 210 232 252 297 326 327 380 398	11
913	Four pairs of vertical strokes enclosing the sign		 140 143 164 179 308	5
914	Angled hat		 196 203 207 411	4

Table 2. (continued)

Modifier	Description	Design	Sign(s) in which it appears	Total Freq.
915	Long angular line		 275 407	4
916	Three pairs of vertical strokes and one single stroke enclosing the sign		 289	1
917	Horn like attachment		 8 9	2
918	Leaf like attachment		 62 295 300 370 376	8
919[@]	Line with attachment to the sign		 25 26 52 183 246 248 269 349 360 413	10
920	Several strokes at the top of the sign		 69 272 399	3
921	Several strokes at the bottom of the sign		 332 337	2

[@] Modifier 919 is generally seen at the left hand side of the basic signs in all cases except when the basic sign is sign number 1  . The left and right directions are as the reader would see them.

3.2.2 Usage of provisional basic signs and modifiers with respect to sites

In Tables 3 and 4 we provide the frequency distribution of provisional basic signs and modifiers respectively, with respect to different sites listed in M77. The provisional basic sign or the modifier is listed in the first column, its design is given in column 2 and columns 3 to 9 give the

frequency of the provisional basic sign or modifier at different sites viz. Mohenjodaro (M), Harappa (H), Lothal (L), Kalibangan (K), Chanhudaro (C), Other Harappan sites (OH) and West Asian (WA) sites. The total frequency of occurrence of the provisional basic sign or modifier is listed in column 10.

Table 4 suggests that the practice of adding modifiers to signs is seen at almost all sites irrespective of the variation of the size of the inscribed corpus from each site.

Table 3. Provisional basic sign usage with sites.

Serial No.	Provisional basic sign	Design	Sites							Total Freq.
			M	H	L	K	C	OH	WA	
			Percentage of total contribution to M77							
			54	33	6	3	2	1	1	100
1	801		0	2	0	0	0	0	0	2
2	802		4	1	0	0	0	0	0	5
3	803*		17	18	0	0	0	1	1	37
4	804		7	1	0	0	0	0	0	8
5	805		1	1	0	1	0	0	0	3
6	806		1	0	0	0	0	0	0	1
7	807		12	4	0	1	0	0	0	17
8	808		1	0	0	0	0	0	0	1
9	809		9	3	0	0	0	0	1	13
10	810		160	60	17	42	4	2	1	246
Total			212	90	17	44	4	3	3	333

Table 4. Modifier usage with sites.

Serial No.	Modifier	Design	Sites							Total Freq.
			M	H	L	K	C	OH	WA	
			Percentage of total contribution to M77							
			54	33	6	3	2	1	1	100
1	901		16	7	0	1	1	0	0	25
2	902		12	2	0	0	0	0	1	15
3	903		170	119	17	3	9	3	2	323
4	904		185	128	20	3	9	2	2	349
5	905		192	69	23	11	11	5	2	313
6	906		74	25	12	2	5	2	0	120
7	907		186	73	21	3	7	5	3	298
8	908		61	24	5	3	3	1	0	97
9	909		0	1	0	0	0	0	0	1
10	910		111	58	18	4	3	2	0	196
11	911		23	13	1	1	2	0	0	40
12	912		105	80	8	7	5	1	4	210
13	913		6	0	0	0	1	0	0	7
14	914		4	1	0	0	0	0	0	5
15	915		43	8	1	0	1	0	0	53

Table 4. (continued)

Serial No.	Modifier	Design	Sites							Total Freq.
			M	H	L	K	C	OH	WA	
			Percentage of total contribution to M77							
			54	33	6	3	2	1	1	100
16	916		4	1	0	0	0	0	0	5
17	917		75	31	5	1	5	0	0	117
18	918		17	4	0	2	3	0	0	26
19	919		54	14	3	1	0	0	0	72
20	920		5	1	1	0	1	0	0	8
21	921		4	4	0	0	0	0	0	8
Total			1347	663	135	42	66	21	14	2288

3.2.3. Types of composite signs

Based on the design elements defined above, the 263 composite signs in M77 can be further classified into following categories: 1) Compound signs and 2) Modified signs.

1) Compound signs (Ligatures)

Compound signs or ligatures are composites of two or more basic signs or provisional basic signs. Out of the 263 composite signs, 149 signs fall into this category. Some examples of compound signs are sign numbers 9, 30, 33 and 372 (Fig. 4).

2) Modified signs

Modified signs are composites of one or more basic signs (or provisional basic signs) with modifiers. Out of the 263 composite signs, 114 signs are modified by one or more of the 21 modifiers listed in Table 2. Some examples of modified signs are sign numbers 60, 61, 65, 139, 164 and 377

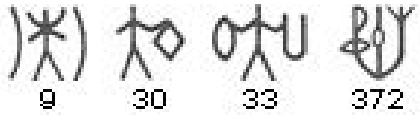


Figure 4. Compound signs.

Figure 5. Modified signs.

(Fig. 5).

4. Decomposition of composite signs

In order to understand the mechanics of the design of Indus signs, we analyse each sign visually and if it is a composite sign it is decomposed into its constituent elements. We do not consider aspects such as co-occurrence of basic signs together in a text or other contextual issues such as associated field symbol, stratigraphy etc. while deciding on the constituent elements of a sign.

The decomposition of some of the composite signs into their constituent elements is shown in Figs. 6 and 7 below. For example, sign number 33 can be decomposed into sign numbers 1, 373 and 328 (Fig. 6), all three being basic signs. Similarly, sign number 61 is made up of sign numbers 59 and 197 with a modifier 906 (Fig. 7). The dataset thus generated for the complete list of 417 signs is then used for further analysis. While performing the analysis we do not assign preference to any order of the constituents and assume that all combinations of the constituents are equally probable.

The decomposition of sign numbers 418 and 419 that are late additions to M77 (see Mahadevan, 1977: 15) are listed in Table 5. In terms of their design, they conform to the general pattern. Hence, it is possible that with increasing corpus size due to new excavations more of such subtly modified signs may continue to be found.

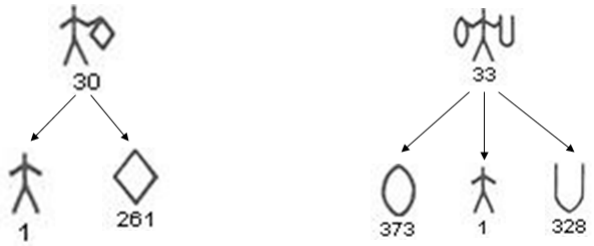


Figure 6. Decomposition of some compound signs: Sign numbers 30 and 33.

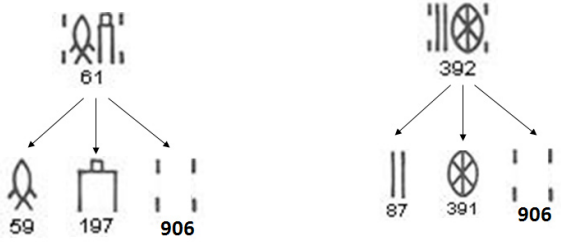


Figure 7. Decomposition of some modified signs: Sign numbers 61 and 392.

Table 5. Analysis of sign numbers 418 and 419 in M77.

Sign No.	Sign appearance	Components				
418	 418	 12	 59	 342	 373	 171
419	 419	 342	'	'	'	'

The classification of 417 signs in the sign list of M77 based on their design is summarised in Fig. 8 below.

4.1 Analysis of compound signs

Out of the 263 signs that have been identified as composite signs, 133 signs are of compound type i.e. made up of two or more of the 154 basic signs. For each of these 133 signs we search whether their constituent basic signs occur in any possible combination in M77. If any combination of

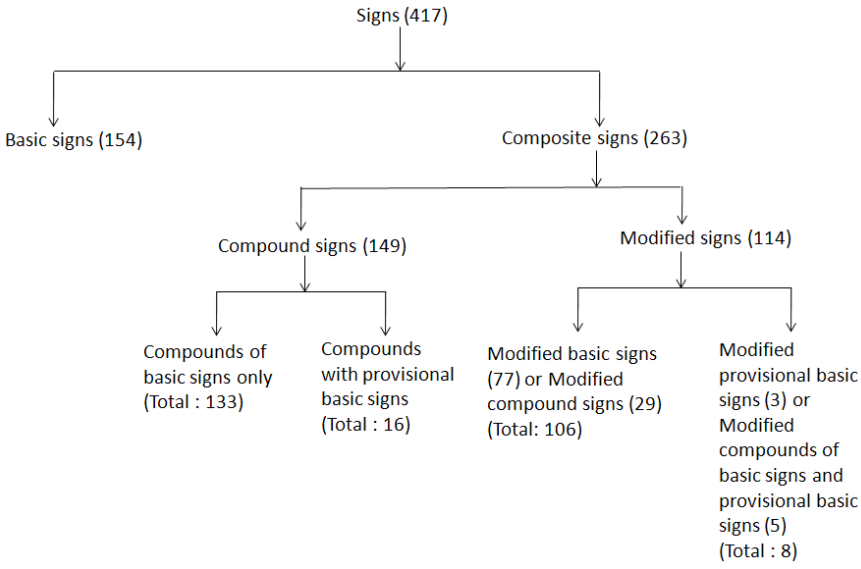


Figure 8. Classification of 417 signs in M77 based on their design.

the constituent basic signs is found in M77, we compare its environment (signs preceding or following) and its pattern of occurrence with that of the compound sign. We find that out of the 133 compound signs, only 28 compound signs have some combination(s) of their constituent basic signs in M77. These 28 cases are listed in Table 6. The summary of the results of this analysis is shown in Fig. 9.

As can be seen from Table 6 and Fig. 9, in most of the cases, the constituent elements in any combination do not share identical environment in comparison to the corresponding compound sign indicating that compound signs are not merely compacted version of the basic sign sequences created for brevity and to save writing space. They seem to have some different function. These signs, compounded with different components, may have stood for some distinct entity or information. The result also justifies the identification of these compound signs and the constituent basic signs as distinct signs in the sign list.

A comparison of basic and composite sign frequency distribution is given in Table 7 and Fig. 10. Several composite signs fall in the high frequency range suggesting that they may have represented some specific entity of common usage.

Out of the 65 signs lying in the frequency range of 50 to 499 in M77,

Table 6. Analysis of compound signs.

Serial No.	Compound sign	Freq.	Components			Sign combination found in M77	Freq.	Comments
1	 15	126	 12	 342		 12 342	8	Sign  48 precedes the compound sign 4 times and the sign combination twice. Compound sign ends texts 94 times and the sign combination ends texts 6 times.
2	 21	2	 1	 162		 1 162  162 1	2 1	No signs common on either side in both the cases.
3	 22	3	 1	 176		 176 1  1 176	6 1	No signs common on either side in both the cases.
4	 38	6	 1	 176		 176 1  1 176	6 1	Compound sign ends texts twice and the sign combination ends texts 6 times. No signs common on either side.
5	 41	20	 1	 134		 134 1	1	No signs common on either side.
6	 141	4	 87	 137		 87 137	1	No signs common on either side.

Table 6. (continued)

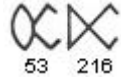
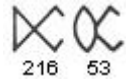
Serial No.	Compound sign	Freq.	Components			Sign combination found in M77	Freq.	Comments
7	 152	1	 150	 162		 162 150	1	Both compound sign and the sign combination begin texts once.
8	 165	1	 162	176		 176 162	4	Compound sign appears solo and the sign combination ends texts twice.
9	 166	1	 162	176		 176 162	4	Compound sign ends text once and the sign combination ends texts twice.
10	 219	11	 216	 53		 53 216  216 53	1 1	No signs common on either side in both the cases.
11	 221	4	 216	 86		 216 86	1	No signs common on either side.
12[#]	 236	1	 267	 180		 180 267  267 180	1 1	No signs common on either side in both the cases.
13	 270	1	 162	 267		 267 162	3	Sign 99 precedes the compound sign and the sign combination once. Both compound sign and the sign combination begins text once.

Table 6. (continued)

Serial No.	Compound sign	Freq.	Components			Sign combination found in M77	Freq.	Comments
14	 273	1	 267	 176	 176	 176 176 267	1	No signs common on either side.
15	 274	1	 267	 176	 176	 176 176 267	1	Both the compound sign and the sign combination end text once.
16	 276	2	 267	 99		 99 267	291	Sign  precedes the compound sign and the sign combination once. Compound sign begins a text once and the sign combination begins texts 260 times. Compound sign ends a text once and the sign combination ends texts 11 times).
17	 282	3	 267	 176		 176 267	1	No signs common on either side.
18	 305	3	 304	 176		 176 304	1	No signs common on either side.
19	 324	2	 373	 373		 373 373	4	No signs common on either side.
20	 350	1	 347	98		 347 98	5	No signs common on either side.

Table 6. (continued)

Serial No.	Compound sign	Freq.	Components			Sign combination found in M77	Freq.	Comments
21	 352	1	 342	 347		 342 347	110	Sign  176 follows the compound sign once and the sign combination 16 times.
						 347 342	1	Sign  176 follows the compound sign and the sign combination once.
22	 356	1	 347	 135		 347 135	5	No signs common on either side.
23	 378	2	 373	 86		 86 373	1	No signs common on either side in both the cases.
						 373 86	1	
24	 387	102	 373	 162		 162 373	2	Compound sign begins texts 30 times and the sign combination begins text twice. Sign  402 follows the compound sign and the sign combination once.
25	 389	134	 373	 169		 373 169	10	No signs common on either side.

Table 6. (continued)

Serial No.	Compound sign	Freq.	Components			Sign combination found in M77	Freq.	Comments
26	 394	2	 342	 373		 373 342	4	No signs common on either side. Compound sign begins a text once and the sign combination begins texts thrice.
						 342 373	4	Sign  176 follows the compound sign once and the sign combination twice.
27	 403	93	 373	 373		 373 373	4	Compound sign begins texts 18 times and the sign combination begins texts twice. Sign  302 follows the compound sign and the sign combination once. Compound sign ends texts thrice and the sign combination ends texts twice.
28	 417	5	 415	 86		 415 86	2	Compound sign appears solo once and the sign combination ends texts twice.

Sign number 236 is erroneously read as  236 in M77 sign list. Its correct version has a round attachment at the top of the inner component (Mahadevan: private communication). We decompose the sign based on the correct version of this sign.

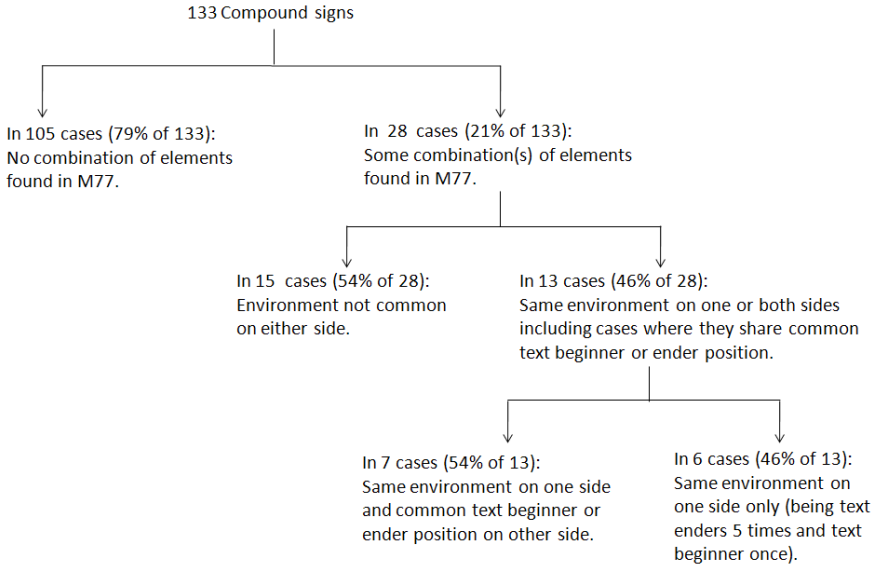


Figure 9. Result of analysis of 133 compound signs.

Table 7. Frequency distribution of basic and composite signs in M77.

Frequency range	No. of signs in M77	Basic signs		Composite signs	
		No. of signs	Percentage of total basic signs	No. of signs	Percentage of total composite signs
1	112	17	11	95	36
2-9	152	48	31	104	40
10-49	86	40	26	46	17
50-99	34	25	16	9	3
100-499	31	22	14	9	3
500-999	1	1	1	0	0
>1000	1	1	1	0	0

47 signs (72% of 65) are basic and remaining 18 signs (28%) are composite (Table 7). However, the balance tilts in favour of composite signs for the less frequent signs in M77. Statistically, out of 112 signs that appear only once in M77 (singletons), only 17 signs (15% of singletons) are basic and 95 signs (85% of singletons) are composite. Hence while high frequency signs tend to be basic, some of the composite signs do have a high frequency of usage.

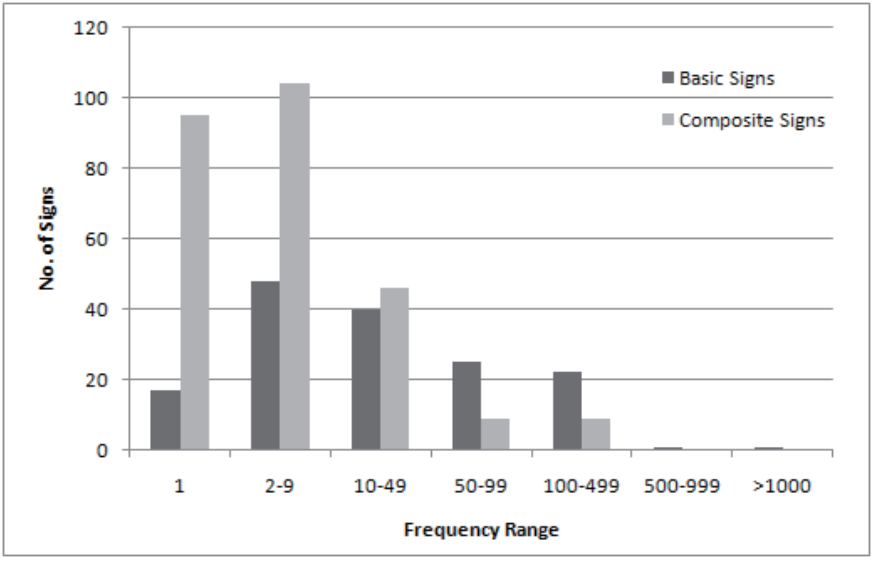


Figure 10. Comparison of basic and composite sign frequency distribution in M77.

5. Effect of modifiers: A case study with modifier 905

We discuss the effect of the modifier 905 on sign number 342 and 328 . Table 8a lists the positional distribution of sign numbers 342 , 343 , 344 and 345 and Table 8b lists the positional distribution of sign numbers 328 , 329 and 330 in M77. Sign number 342 is the most frequent sign in M77 and is also the most frequent text ender. Tables 8a and 8b suggest that as the number of modifier 905 increases in sign numbers 342 and 328 their position in the Indus texts moves towards the beginning of the texts.

An indication of the possible logic is also seen in the usage of another modifier 906 (Hunter, 1934; Mahadevan, 1986). They have suggested that 906 seems to replace frequent text ender signs such as sign number 211 or 342 indicating that addition of the modifier 906 to a sign makes it a *pseudo ender*. The effect of all modifiers is out of scope of the present paper and will be explored in detail in our forthcoming work.

Table 8a. Effect of modifier 905 on sign number 342 .

Sign No.	Sign Image	Frequency	Solo (%)	Initial (%)	Medial (%)	Final (%)
342		1395	0.22	0.07	30.11	**69.61**
343		177	0.00	11.30	**85.88**	2.82
344		35	8.57	11.43	**74.29**	5.71
345		51	0.00	**49.02**	47.06	3.92

Table 8b. Effect of modifier 905 on sign number 328 .

Sign No.	Sign Image	Frequency	Solo (%)	Initial (%)	Medial (%)	Final (%)
328		323	0.31	5.88	8.98	**84.83**
329		7	0.00	28.57	**71.43**	0.00
330		8	0.00	**62.50**	37.50	0.00

Note that sign number 328 does not occur with single occurrence of the modifier 905. That is, there is no equivalent of sign number 343 in the 328 series.

6. Some interesting aspects of Indus signs

In this section, we explore some additional features of Indus signs. We begin with the statistics of symmetry in the design of Indus signs.

6.1 Symmetries in the design of Indus signs

The basic statistics of symmetry in the design of the 417 signs of M77 are given in Table 9. We investigate mirror symmetries in vertical and horizontal planes around the centre of the sign image. Our study is based on the normalised sign list and while implementing these designs in real

Table 9. Basic symmetries in Indus sign design.

Serial No.	Type of symmetry	No. of basic signs	Percentage of total basic signs	No. of composite signs	Percentage of total composite signs	Total no. of signs	Percentage of total signs in M77
1	Vertical symmetry	93	60	127	48	220	53
2	Horizontal symmetry	58	38	47	18	105	25
3	Both horizontal and vertical symmetry	40	26	30	11	70	17
4	No symmetry	43	28	119	45	162	39

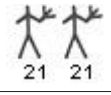
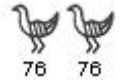
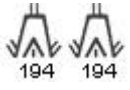
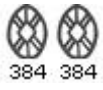
life these symmetries would have been skewed by Indus writers. Kenoyer (2006) illustrates how the quality of sign engraving improves significantly during the mature phase of the Indus valley civilisation. We therefore assume that the normalised sign list is the idealised sign list that the Indus people would have used and ignore the aberrations due to the frailty of human hand. With these approximations we find that out of 417 signs, 255 signs have either vertical symmetry, or horizontal symmetry or both and 162 signs have no symmetry at all.

6.2 Doubling and repetition

Several signs in M77 appear in pairs where the same sign repeats itself. There are 30 such pairs in M77. The list of pairs with self-repeating signs in M77 is given in Table 10. Similarly, there are six combinations in M77 with same sign repeating more than twice and they are listed in Table 11.

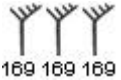
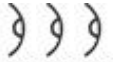
Some of the combinations listed in Table 11 may have served numerical function. An example suggestive of such a function is the text  that occurs twice in M77 (as text number 2698 in Mohenjodaro and text number 7002 in Lothal) and the text  that occurs once in Mohenjodaro (text number 2322 in M77). It is interesting that the three signs (sign numbers 328 , 162 and 169) that have multiple

Table 10. Sign pairs with self-repeating signs (sorted by frequency).

Serial No.	Sign pair	Freq.	Serial No.	Sign pair	Freq.	Serial No.	Sign pair	Freq.
1	 245 245	70	11	 89 89	7	21	 21 21	1
2	 1 1	14	12	 293293	4	22	 76 76	1
3	 153 153	10	13	 373 373	4	23	 104 104	1
4	 328 328	10	14	 323 323	3	24	 212 212	1
5	 121 121	9	15	 54 54	2	25	 252 252	1
6	 176 176	9	16	 127 127	2	26	 290 290	1
7	 375 375	9	17	 169 169	2	27	 307 307	1
8	 391 391	9	18	 194 194	2	28	 324 324	1
9	 162 162	8	19	 237 237	2	29	 342 342	1
10	 59 59	7	20	 384 384	2	30	 389 389	1

appearances in the first four rows of Table 11 also appear with some of the vertical stroke signs such as sign numbers 89  , 102 and 104 , but the two signs in the last two rows of Table 11 do not appear with such vertical stroke signs.

Table 11. Sign combinations with same sign repeating more than twice.

Serial No.	Sign combination	Frequency in M77
1	 328 328 328	3
2	 162 162 162 162	2
3	 162 162 162	1
4	 169 169 169	1
5	 237 237 237	1
6	 293293293	1

6.3 Mirroring

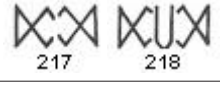
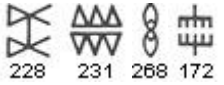
Some signs in the sign list of M77 show the effect of mirror reflection in their design. The signs showing mirror reflection in their design along the vertical and horizontal axes are listed in Table 12.

It is to be noted that the sign components of the signs listed in Table 12 never occur as a sign combination in M77 (except for the sign number 176  that appears as a pair nine times in M77) suggesting that sign compounding is not meant for making writing concise. Sign number 346  is included in this table under the assumption that the inverted 'V' shaped component with the three short vertical strokes at the bottom in its design is an inversion of sign number 345  .

6.4 Signs with different configuration of identical components

In Table 13 we list sets of compound signs that have the identical

Table 12. Different mirroring configurations in Indus sign design.

Serial No.	Configuration	Signs
1	Mirror reflection along vertical axis	 217 218
2	Mirror reflection along horizontal axis	 228 231 268 172
3	Other configurations	 133 346 382

constituents but are combined in different ways to create different compound signs.

For each case listed in Table 13, we have analysed the environment of each compound sign sharing identical components in different design configuration. We find that in *none* of the first ten cases do the compound signs share common environment except for their appearance at beginner or ender positions. In case of compound signs in row 11 (Table 13), there is a broad similarity in the signs that appear in the neighbourhood of the two compound signs. However, even in this case the environment is not identical. This reinforces our earlier suggestion that compound signs are not meant for brevity but seem to have some different function.

7. Discussion

The sign list of M77 provides a broad overview of the design of the Indus signs. Here, we identify three kinds of elements that go in the making of Indus signs. They are: basic signs (154), provisional basic signs (10) and modifiers (21). The *basic signs* are relatively simple in their design but the *composite signs* are often a combination of two or more elements defined above. The designs of Indus signs imply logic and creativity in their makeup. The signs are often merged so that they retain aesthetic quality. About 60% of the signs conform to either vertical or horizontal symmetry (Table 9). While the visual form of some of the signs can be directly associated with familiar natural or artificial entities, a large number of

Table 13. Signs with same components but different design configuration.

Serial No.	Signs with same components but different configuration		Components			
1	 22	 38	 1	 176		
2	 40	 43	 1	 237		
3	 63	 64	 59	 78	 287	 299
4	 93	 94	86	 87		
5	 165	 166	 162	 176		
6	 172	 173	 171	 171		
7	 228	 309	 98	 225	 225	
8	 264	 284	 149	 261		
9	 273	 274	 267	 176	 176	
10	324	403	373	373		
11	375	386	98	373		

signs are characterised by a high level of abstraction in their design.

In order to estimate the complexity in writing these signs, we analyse the structure of signs based on the number of strokes required to write a

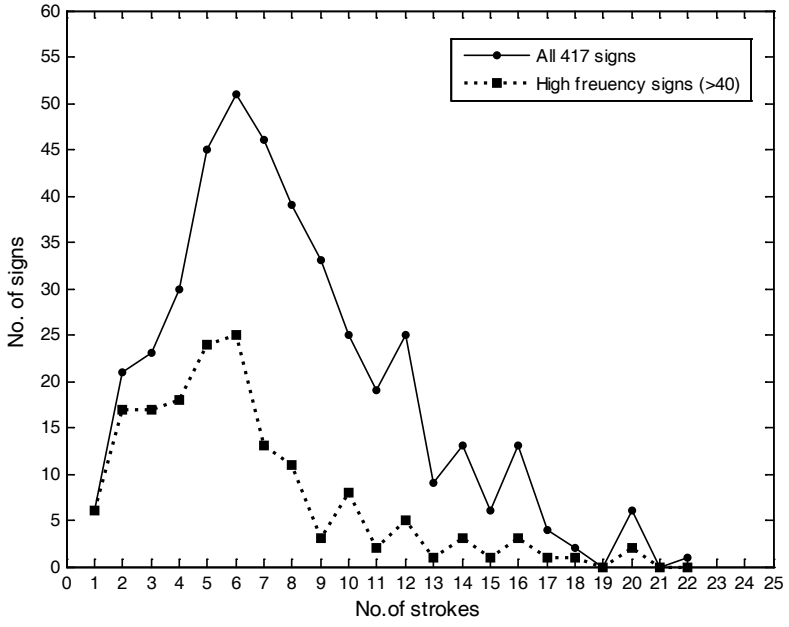


Figure 11. Number of strokes in the design of Indus signs.

sign and correlate it to the number of signs. In Fig. 11 we plot the number of strokes required to write a sign against the number of signs. As can be seen from Fig. 11, most signs require between five to eight strokes while writing, when all the signs are considered. The average number of strokes required while writing high frequency signs is five strokes. However, the tail of the distribution of number of strokes per sign is fairly long and signs requiring more than fifteen strokes are seen even amongst high frequency signs. This suggests that ease of writing does not seem to be a criterion in designing and usage of signs.

Changizi et al. (2006) have shown that the design configuration across various types of visual signs seem to have been deeply influenced by the shapes that the designers of the signs encounter in their daily lives rather than ease of writing. This can also be seen in the context of Indus script suggesting that the writers were very keen on conveying their ideas or information unambiguously using the shapes found in their environment rather than placing emphasis on ease of writing. Moreover, several multi-stroked signs are ligatures of various basic signs. This gives an impression

that some form of shorthand was being tried. However, our analysis shows that this does not seem to be the case and this makes the problem of the Indus writing even more complex.

The compound signs are not mere compacted forms of the constituent basic signs as the constituents rarely appear as sign sequences in the Indus texts. Even in the rare cases when the components of a compound sign appear in some combination(s) in M77, the environment of the compound sign and that of the sequence of its constituent basic signs is very different (Table 6). The study of the environment of compound signs and the combination(s) of their constituents provides important clues. It is clear from the study that:

a) Almost always, the environment of a compound sign is different from the combination of the basic signs that were merged together.
b) In almost 80% of the cases (105 out of the 133 cases), the original constituents of the compound sign do not even appear as a combination.

This suggests that there is no obvious rule as to how a compound sign should be decomposed. Hence economising the usage of writing space does not seem to be the rationale in creating compound signs. While the basic signs dominate the high frequency range, the composites dominate the low frequency range (frequency less than 10, see Table 7). The signs use techniques of doubling (Table 10) or repeating themselves more than twice at times (Table 11) as well as reflection along horizontal or vertical axis in their design (Table 12). At times two compound signs differ only in the configuration of identical components (Table 13). Detailed analysis of such compound signs reveals that the the placement or location of the constituents in the design of compound signs plays an important role in the design of Indus signs (Table 13).

There are 21 modifiers, sign elements that do not have independent existence but which clearly seem to add value to a significant number of signs across the sign list at least in some cases (Tables 8a, 8b). This suggests that the modifiers add a value that can affect the utility of the signs. This is apparent from the data presented in Tables 8a and 8b, where addition of the modifier 905 to sign numbers 342  and 328  significantly changes their position in the texts, but the variation that the modifier produces, appears similar.

The signs of the Indus script seem to incorporate techniques in their design that were used in several ancient writing systems to maximise the usage of limited number of signs. It seems to incorporate sign compounding as seen in Chinese writing (Bottéro, 2004), conflation of signs as done in Mayan glyphs (Coe, 1992) and it also seems to have signs functioning as determinatives as was the practice in Egyptian hieroglyphs (see Baines, 2004; Parpola, 1994). In later scripts in India merger of signs is used to combine vowels into consonants. However, we do not wish to hazard a guess on the connection or otherwise of Indus script with various writing systems.

8. Conclusions

We have studied the general makeup of the Indus signs by identifying their design elements based on the normalised sign list of M77. The principal design elements of the Indus signs that are identified by the study are basic signs (154), provisional basic signs (10) and modifiers (21). The number of basic signs is consistent with Parpola (1994) but far higher than the reduction to 20 basic signs proposed by Rao (1982). The reason for this discrepancy is that our classification of the 417 signs into basic and composite signs is *based solely on the design of a sign*. We do not assign meaning or value to what the sign may stand for and therefore the analysis is *independent of any assumption about its contents*.

Composite signs are further classified into compound signs and modified signs. We analyse the occurrence pattern of all compound signs and compare it with all possible sequences of their constituent basic signs in M77. Our results suggest that compound signs are not meant for brevity and saving writing space but seem to have some other function. Our major conclusions are listed in Table 14.

The signs have been designed with care and combining signs with other signs or modifiers seems to have been a practice known to all sites. Indus people seem to employ several interesting techniques such as using a set of modifiers to modify the basic sign and sign compounding. We also find instances of doubling and mirroring in the design of Indus signs. The usage of a similar set of constituent basic signs in different configuration in

Table 14. Major Conclusions.

Serial No.	Analysis	Table/ Fig. No.	Results and conclusions
1.	Design elements	Tables 1, 2	Basic signs: 154, Provisional basic signs: 10, Modifiers : 21
2.	Classification of signs	Fig. 8	Basic signs: 154 and Composite signs: 263 Composite signs are further divided into: Compound signs: 149 and Modified signs: 114
3.	Site-wise distribution of provisional basic signs and modifiers	Tables 3, 4	Modification of signs was practiced at all sites.
4.	Analysis of compound signs	Table 6, Fig. 9	The environment of compound signs and their constituent basic signs (in various combinations) are different suggesting that they had different function than the sequence of their components.
5.	Frequency distribution of basic vs. composite signs	Table 7, Fig. 10	Basic signs often dominate the high frequency range. Composite signs dominate the low frequency range.
6.	Case study of modifier 905	Tables 8a, 8b	Modifier 905 shows similar effect on sign numbers 342 and 328.
7.	Symmetry in sign design	Table 9	More than 60% of signs have some form of symmetry.
8.	Doubling and repetition	Tables 10, 11	Thirty signs appear in doublets with themselves and 6 signs show repetition more than twice.
9.	Mirroring	Table 12	Six signs show complete vertical or horizontal mirror reflection effect in design while three signs have other interesting configuration.
10.	Sign with identical components but different configuration	Table 13	In eleven cases the same set of basic signs combine in two different ways but their environment is different suggesting that the location of components within the compound sign is important.
11.	Analysis of number of strokes required to write a sign	Fig. 11	On an average about eight strokes are required to write an Indus sign but high frequency signs which require an average of five strokes,span almost the entire range of stroke requirements suggesting that ease of writing does not seem to be a criterion in designing and usage of signs.

compound signs seem to convey different information. The designers of the Indus signs also placed a special emphasis on symmetry and there seems to be an underlying effort to retain the overall aesthetic value of Indus signs.

All this makes it clear that like all other writing systems, Indus writing is an intellectual exercise of great significance and a lot of thought, planning and utility issues have been taken into consideration while designing these signs. The Indus civilisation was spread over an area of about a million square kilometres and yet, the sign list over the entire civilisation seems to be the same indicating that the signs, their meaning and their usage were agreed upon by people with large physical separation. This arrangement worked satisfactorily for about 700 years. Hence the understanding of Indus signs and their meaning must have been robust and yet versatile and easy to use. This is also reaffirmed by the fact that Indus script has been found on seals discovered in West Asia with a different grammar (Rao et al., 2009b). The usage of modifiers at almost all sites (Table 4) suggests that the manner of sign modification was universally agreed over the entire area of the Indus valley civilisation and was not intended for a small group of people.

Acknowledgement

We wish to acknowledge the support of Jamsetji Tata Trust for this work. We also wish to thank Dr. Iravatham Mahadevan for his insightful guidance. We would like to thank Mr. Hrishikesh Joglekar for his help in designing the software used for generating text images. We are grateful to Harappa Archaeological Research Project, J.M. Kenoyer and Harappa.com for their kind permission to use the picture of an Indus seal in the paper. We wish to thank the two unknown referees whose valuable suggestions have greatly improved the paper.

References

Agrawal, D. P. (2007), *The Indus Civilization: An Interdisciplinary perspective.* New Delhi: Aryan Books International.

Baines, J. (2004), The earliest Egyptian writing: development, context,

purpose. In: Houston, S. D. (Ed) *The First Writing: Script Invention as History and Process.* Cambridge University Press, Cambridge, 150-189.

Bottéro, F. (2004), Writing on shell and bone in Shang China. In *The First Writing: Script Invention as History and Process,* Houston, S.D. (ed.). Cambridge: Cambridge University Press: 250-261.

Changizi, Mark A., Zhang, Qiong, Ye, Hao and Shimojo, Shinsuke (2006), The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes, *The American Naturalist*, Vol. 167, No.5.

CISI = *Corpus of Indus Seals and Inscriptions. 1: Collections in India,* edited by Jagat Pati Joshi and Asko Parpola, 1987. 2: *Collections in Pakistan,* edited by Sayid Ghulam Mustafa Shah and Asko Parpola, 1991. 3: *New material, untraced objects, and collections outside India and Pakistan*, edited by Asko Parpola, B. M. Pande and Petteri Koskikallio, Part 1: *Mohenjodaro and Harappa*, in collaboration with Richard H. Meadow and Jonathan Mark Kenoyer, 2010, (Annales Academiae Scientiarum Fennicae B 239-241.) Helsinki: Suomalainen Tiedeakatemia.

Coe, M. D. (1992), *Breaking the Maya Code*. New York: Thames and Hudson.

Damerow, P. (2006), The Origins of Writing as a Problem of Historical Epistemology, *Cuneiform Digital Library Journal 2006:1.*

Daniels, P. T. and Bright, W (eds.) (1996), *The World's Writing Systems,* New York: Oxford University Press.

Houston, S. D. (ed.) (2004), *The First Writing: Script Invention as History and Process,* Cambridge: Cambridge University Press.

Hunter, G.R. (1934), *The Script of Harappa and Mohenjodaro and its Connection with Other Scripts*, London: Kegan Paul, Trench, Trubner & Co. Ltd.

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, in *Proceedings of the Pre-symposium of RIHN and 7th ESCA Harvard - Kyoto Roundtable*, ed. Osada Toshiki.

Mahadevan, I. (1977), *The Indus Script: Texts, Concordance and Tables*, Memoirs of the Archaeological Survey of India No. 77.

Mahadevan, I. (1986), Towards a grammar of the Indus texts: 'Intelligible to the eye, if not to the ears', *Tamil Civilization* 4 (3 & 4): 15-30.

Mahadevan, I. (1998), Phonetic value of the arrow sign in the Indus script, *Journal of the Institute of Asian Studies* 15 (2): 69-74.

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.

Mahadevan, I. (2006), Agricultural terms in the Indus script, *Journal of Tamil Studies*, Vol. 70, pp. 64-76.

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*, no. 50: 28-66. Tokyo: The Tôhô Gakkai.

Parpola, A. (2009), 'Hind Leg' + 'Fish': Towards Further Understanding of the Indus Script, *Scripta*, Vol. 1, pp. 37-76.

Pope, M. (1999), *The Story of Decipherment: From Egyptian hieroglyphs to Maya script*. London: Thames and Hudson.

Possehl, G. L. (1996), *Indus Age: The writing System*. New Delhi: Oxford IBH Publishing Co. Pvt. Ltd.

Possehl, G. L. (2002*), The Indus Civilization: A contemporary perspective.* New Delhi: Vistaar Publications.

Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., 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., Mahadevan, I. (2009b), A Markov model of the Indus script, *Proceedings of the National Academy of Sciences,* 106, 13685.

Rao, R. P. N., Yadav, N., Vahia, M. N., Joglekar, H., Adhikari, R., Mahadevan, I. (2010), Entropy, the Indus Script, and Language: A Reply to R. Sproat, *Computational Linguistics* 36 (4), pp. 795-805.

Rao, S. R. (1982), *The Decipherment of the Indus Script*. Bombay: Asia Publishing.

Renfrew, C. (1989), *Archaeology and Language*. London: Penguin Books.

Robinson, Andrew (2009), *Lost Languages, The enigma of the world's undeciphered scripts*. London: Thames and Hudson Ltd.

Siromoney, G. and Haq, A. (1988), Segmentation of Indus Texts: A dynamic programming approach, *Computers and the Humanities,* Vol. 22, pp. 11-21.

Subbarayappa, B. V. (1997), *Indus Script: Its nature and structure*. Madras: New Era Publications.

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., Vahia, M. N., Mahadevan, I., Joglekar, H. (2008a), A statistical approach for pattern search in Indus writing, *International Journal of Dravidian Linguistics*, XXXVII Number 1, pp. 39-52, http://www.harappa.com/script/tatawriting/tata-indus-script-research.html.

Yadav, N., Vahia, M. N., Mahadevan, I., Joglekar, H. (2008b), Segmentation of Indus Texts, *International Journal of Dravidian Linguistics*, XXXVII, Number 1, pp. 53-72, http://www.harappa.com/script/tatawriting/tata-indus-script-research.html.

Yadav, N., Joglekar, H., Rao R. P. N, Vahia, M. N., Adhikari, R., Mahadevan, I. (2010), Statistical Analysis of the Indus Script using n-grams, 2010, *PLoS One*, Vol. 5, Issue 3. http://dx.plos.org/10.1371/journal.pone.0009506.