Why the Indus Script WAS true writing and why a larger corpus of texts existed in the Indus Valley civilization: Simple proof addressed to mainstream researchers & archaeologists

Sujay Rao Mandavilli

This paper is meant to read together with the paper 'The reconfirmation and reinforcement of the Indus script thesis: a logical assessment and inquiry as to the elusive and enigmatic nature of this script', which was published in the ICFAI Journal of History and Culture in January 2011. In the aforementioned paper, we had clearly shown that the Indus script used in the Indus Valley civilization which flourished from 2600 BC to 1900 BC, was a logo-syllabic script. In this paper, we show that the case for the lost manuscript hypothesis has never been stronger than it has been in the past one decade. This hypothesis has had many adherents even in the West even earlier when very little of the Indus had been excavated, but few will now deny that no other scenario is likely. This hypothesis was earlier based on hunches, now its adherents can base it on science and valid epistemology. The Indus Valley Civilization has always amazed legions of archeologists since the 1920's and has been taught to students all over the world: it can now take its pride of place among old world civilizations. This paper also introduces Logo-syllabic thesis B as opposed to the older logo-syllabic thesis A and lays bare the differences between the two. We insist that only approaches such as those detailed in this paper can be applied for the study of the Indus script given the low quantum of archaeological data in relation to the total known size of the IVC. (This ratio is the lowest for all known civilizations). This paper is meant to be read after 'Syncretism and Acculturation in Ancient India; a new nine phase acculturation model explaining the process of transfer of power from the Harappans to the Indo-Aryans' which was published in two parts in the peer-reviewed ICFAI journal of History and culture in January 2009 and January 2010. This paper detailed methods to reconstruct the languages spoken in the IVC.

In the earlier paper, 'The reconfirmation and reinforcement of the Indus script thesis: a logical assessment as to the elusive and enigmatic nature of this script', which was published in the ICFAI Journal of History and Culture in January 2011[1], we had shown very logically, and in a manner that would have appealed to the skeptic and the layman alike, that the Indus script was a fully evolved logo-syllabic script. In this paper, we take our earlier conclusions a step further and show irrefutably that a larger corpus of texts existed in the Indus. The Lost Manuscript hypothesis has had many adherents even in the West even decades earlier when very little of the Indus had been excavated, but few will now deny that no other scenario is likely. This hypothesis was earlier based on hunches, now its adherents can base it on science and valid epistemology. The Indus Valley Civilization has always amazed legions of archeologists since the 1920's and has been taught to students all over the world: it can now take its pride of place among old world civilizations. Archeologists should always keep this in mind: a longer corpus of texts may well turn up someday if inscribed on non-perishable materials.

The reasons why a large corpus of texts definitely existed, either on non-perishable or perishable materials, and can be very reliably inferred are given below:

Reason # 1 The Dholavira Signboard

Fig 1. This signboard displays a high rate of sign repetition, more than Linear-Elamite inscriptions even. The fact that the order of signs could have played a role in determining the meaning cannot be overlooked , and possible evidence for this can come from the fact that the spoked wheel occurs in the second and third position in the signboard (also the seventh and the tenth). This glyph may have encoded speech as this signboard may have represented the name of this place. Note how this feature may have been put to use in land-locked and far-from Mesopotamia Dholavira for ornamental purposes, pointing out to a greater use of this feature, because this relatively difficult to use option would have been imported only if it had served a greater purpose. This automatically implies the existence of longer texts. This also proves one thing: speech-encoding could be interpreted by a relatively large number of elites.

One of the most significant discoveries in the Indus was made in the Indus city of Dholavira, now in Gujarat in India. It was a large signboard with ten characters discovered in 1999. Each sign measured about 37 centimeters high and the board on which letters were inscribed was about 3 meters long, making it very large by Indus standards. Some further facts about the Dholavira signboard are presented below so as to reinforce our case.

---

[1] The Reconfirmation and reinforcement of the Indus script thesis: A logical assessment as to the elusive and enigmatic nature of this script Mandavilli, Sujay Rao ICFAI Journal of History and Culture January 2011

- It is one of the most famous of Harappan inscriptions and has become instantly famous ever since it was discovered in 1999. It also shows convincingly that the Indus script was put to a wide variety of uses.
- It was very large in size, and even giant if compared to the much smaller seals and measured 3 meters or so in length.
- It was located in Far-from-Mesopotamia Dholavira, and in one of the furthest sites from Mesopotamia. It was not even the largest city in the IVC.
- It hung over the citadel there, and may have been a prominent landmark of the citadel. The fact that it was hung over the citadel meant it was meant to be read by large sections of society and at least by elites.
- It must have represented the name of the place and must have been closely tied to speech: note the sign repetition. The rate of sign repetition is much higher than most Linear-Elamite inscriptions. The spoked wheel sign repeats four times.
- Even Rongorongo glyphs, though unquestionably longer, do not display this kind of sign repetition. There is no overwhelming evidence that Rongorongo branched off from a script which had reached the linguistic stage, either.
- The sign which was used as a determinative was a very common Indus sign. I.e. the spoked wheel sign.
- The sign used as a determinative appears to have been also similar to determinatives in other writing systems such as Linear Elamite but not exactly identical to them. This feature is not available in Proto-Elamite. On the other hand, Linear Elamite does not appear until around 2100 BC i.e. well after the Indus script came into widespread use or closer to the end of the IVC. While the relationship between Proto and Linear Elamite is yet to be established (there are several schools of thought, and most consider the link to be tenuous)[2], it would strongly imply that the Indus script was logo-syllabic.
- The glyph suggests that the determinative could be used either before or after another common Indus sigh probably varying its meaning.
- The other signs with which the determinative was used were also common Indus signs.
- The Indus script was also related to Proto-Elamite but was developed after the scripts of West Asia had become logo-syllabic which means it probably had a linguistic component.
- It was put to the most frivolous use.
- Few scholars will now dispute the fact that the Indus script was a logo-syllabic script on the basis of this evidence. Few scholars will deny the fact that speech encoding was one of the major functions of the Indus script and had this feature had reached a very precocious maturity. On the basis of this evidence, we can even conclude that this appears to be a fully developed logo-syllabic script. Human learned to code sound through the rebus principle and acrophony and the transition from logographic to logo-syllabic scripts was often long drawn and messy. This glyph however, appears to represent a logo-syllabic script in a mature form, perfectly adopted for use in the Indus. One would only hope more such inscriptions were found, but to state that the discovery of this inscription was probably a timely godsend would be a gross understatement and would only make our case very strong, in a manner no other glyph perhaps can. The reasons would not have been very far to seek: the Indus script would have had the benefit of being introduced later than the scripts of Egypt or Mesopotamia and all the developmental effort was probably carried out there. The Harappans were spared the trouble of experimentation and imported the script after it had reached maturity.
- This inscription was apparently more closely tied to speech than most Linear-Elamite inscriptions. Speech encoding would have been a prized possession: no one would have used it just for a decorative signboard at far-from-Mesopotamia Dholavira. Why would a man who had inscribed this, done so

    o If nobody else could read it?
    o If it was hung so prominently, at least elites could certainly have read it

---

[2] The archaeology of Elam: Formation and Transformation of the Ancient Iranian state D.T.Potts Cambridge University press, 1999

- o if elites had been able to read it, they would have certainly written in it, otherwise they wouldn't have been trained in speech encoding to begin with.
- o Why would he have learnt to encode speech only to inscribe this signboard? This automatically implies the existence of longer texts. It also shows that the Indus elites used more complex forms of communication.
- This also proves that speech encoding could be interpreted by a large number of elites
- If they had learnt to interpret speech encoding, why would they have wasted their time acquiring these skills only to interpret this signboard? What other proof does one need for a lost corpus?
- Even if we assume that speech-encoding was added in Mature Harappan 3B, this logic would still hold good. This is of course a much less likely scenario.
- This logic is already accepted by mainstream Indus archaeologists as a precursor to the existence of longer texts
- On the basis of this glyph, we can even make a very, very early attempt to ascertain the signs that were meant to be used with determinatives and signs that were non-linguistic: no animated signs appear in this glyph. In addition, we know that the other signs in this glyph are abstract. They are less complex meaning it was designed with longer texts in mind.
- The other signs that appear with the supposed determinative appear to be abstract and simple lending weight to the theory that the Indus script was designed with longer texts in mind.
- This glyph was found in Dholavira, the fifth largest city of the Indus, and was far away from Mesopotamia. If true writing had existed here, it would have most certainly existed in major Indus cities, given the fact that a small portion of the Indus has been discovered.

British archaeologist Jane McIntosh states (her conclusions very strongly reinforce the conclusions presented here. Let us also recall the 1:x ratio we stated earlier in the paper):

"Farmer also draws attention to the absence of long Harappan inscriptions on potsherds. If the Harappan signs were a script, he contends, this absence would make it unique among the scripts of literate cultures, who all used potsherds often like scrap paper. This need only, imply however, that the Harappans had other media that were easier to scribble on, such as cotton cloth or wooden boards, or that the writing medium was not well suited for use on sherds. Likewise the absence of long monumental inscriptions seems significant to Farmer, but the Harappans did not create monumental art or architecture on which such inscriptions might have been written; the nearest they came to this is the Dholavira signboard, which is quite possibly the tip of an iceberg of a now vanished public inscriptions. He (Farmer) also considers that the proportion of singleton and rare signs is unusually high; other scholars such as Parpola (2005) demonstrate that this is not so, since in general logo-syllabic scripts contain a small corpus of frequently used signs and a large number of much less common ones. Moreover, new signs are continuously added, even when the writing system is a fully developed one, something Farmer also denies. Statistically the Harappan script does not differ significantly in its sign proportions from other logographic scripts. A further point regarding the singletons is that Wells (n.d.) has demonstrated that many are variants or ligatures of basic signs, rather than completely different signs; again, this is something to be expected in a genuine script.

"Perhaps more significantly, the brevity of the majority of the Harappan texts (four to five signs on average) makes it less likely that signs would repeat within them than it is in the longer texts with which Farmer compares them (McIntosh 2008, p. 374). Farmer's arguments fail to account convincingly for the structural regularities that analyses have revealed in the use of the Harappan signs; these seem strongly to support the hypothesis that the Harappan signs represent a writing system. The theory put forward by Farmer and his collaborators has not been widely accepted, but it has been valuable in compelling scholars to look afresh at their assumptions about the script and in provoking a stimulating debate from which a deeper understanding of the script should emerge (McIntosh 2008, p. 374).[3]

---

[3] McIntosh, Jane R. (2008). The ancient Indus valley: new perspectives. Santa Barbara, California: ABC CLIO. ISBN: 978-1-57607-907-2 (hard copy; alk paper)

The Dholavira signboard would itself, constitute, in the eyes of most, an almost irrefutable proof that a lost corpus existed, and views of mainstream western scholars were provided even in our earlier paper in support of this very obvious observation, but let us now attempt to take our hypothesis much further from our previous papers.

Reason # 2 Does not conflict with the history of Writing systems, writing materials, implements and the availability of materials in the IVC. All other points raised by Farmer have been refuted in our paper

The availability of longer texts does not in any way conflict with the history of Writing systems, writing materials, implements and the availability of materials in the IVC. All other points raised by Farmer et al have been refuted in our previous paper dealing with the Indus script in addition to the refutations provided by Asko Parpola, Bryan Wells and several other scholars over the past several years, both Western and Indian. [4] This should seek to drive home the point that the Non-script thesis is a complete non-starter. However, Farmer has been correct in questioning the brevity of the inscriptions, and the reasons for this were provided in our paper. As noted, the Indus script was clearly related to the scripts of the Middle East and was developed only after the script there had morphed into logo-syllabic scripts. Inscriptions in scripts such as Etruscan (Alphabetic), Mayan Hieroglyphs (Logo-syllabic) and Proto-Canaanite (Alphabetic) were known to have been short but longer inscriptions were unearthed subsequently. In very early days, people did not know whether Egyptian hieroglyphs were true writing either. There are many cases where writing thought to be proto-writing turned out to be true writing. In the case we can say given the small amount of data available that there is absolutely no chance that longer writing won't turn up in the IVC. The use of the determinative in the Dholavira signboard shows this too. (The use of Proto-Elamite ended in 2700 BC. Linear Elamite did not appear until 2100 BC. The Indus script was developed in 2600 BC and does not relate to the earlier very rudimentary writing of the Indus Valley (please refer our earlier paper). 35 signs of the Indus script can be compared to the Proto-Elamite script. It would also be illuminating to compare the Proto-Elamite script, the Indus script & the Linear Elamite script on a sign by sign basis, and this would be an interesting exercise to undertake.)

We have also covered the transformation of Harappan India to Post-Harappan India in greater detail in our previous papers and have also dealt with literacy in Post-Harappan India. The Indus script did survive in Post-Harappan India, but its usage had reduced as superior writing systems were available.

None of the so-called symbol systems would pass these tests convincingly and there in lies a huge and fundamental difference: can anyone say all that we have said in our earlier paper and in this one about the Indus script about Vinca symbols and Pictish stones?

Reason # 3 The Hypothesis that the presence of many rare signs proves the instability of the Indus script is not only just statistically unreliable at this stage: it is based on a fundamentally wrong premise. Sign distributions can also be used to predict the existence of a lost corpus of inscriptions

We have discussed this hypothesis in detail in our previous paper dealing with the Indus script. This reasoning is wholly invalid for several reasons (all of which were discussed in our paper), the chief of these being:

- This hypothesis is statically unreliable at this stage, and only 5% of the IVC has been excavated.
- This hypothesis also does not take into account the fact signs had to be read in conjunction with other signs like in the contemporary Proto-Elamite script.

---

[4] *The Indus script: A positional –statistical approach, Michael Korvink, 2007*

- The number of rare signs does not vary significantly from contemporary scripts such as Proto-Elamite and a comparative assessment is presented in this paper.
- There are many reasons to adduce that the Indus-script was a writing system, and the idea that the Indus script wasn't even a writing system would fly in the face of all research done by many mainstream scholars over the decades.
- If the Indus script had a linguistic component as admitted by Farmer et all themselves, the rare sign hypothesis ceases to be of any relevance.
- Our assessment: it was a stable system, fully imported and was designed for both non-linguistic and much longer texts and did not change because it was fully stable.

S. No	Type of script	Rare signs	Total number of signs	Is this a characteristic of the Indus script?
1	Logo-syllabic scripts	High	High	Yes
2	Logo-syllabic scripts where the percentage of Word signs out of the total number of signs is high.	Higher	Higher	Yes
3	Logo-syllabic scripts where the percentage of Word signs out of total number of signs is high and where order of signs, position of signs and sign combinations are important in determining the meaning as observed in Proto-Elamite.	Even higher	Even higher	Yes
4	Logo-syllabic scripts where the percentage of Word signs out of total number of signs is high and where order of signs, position of signs and sign combinations are important like proto-Elamite and the script is used over a wide area over a long period and by many people.	Much higher	Much higher	Yes

The nature of a Civilization and its bearing on the type of inscriptions

The percentage of long inscriptions out of the total number of inscriptions will vary depending on the type of the civilization. For civilizations like they IVC, they will be very low. This certainly reinforces the idea that longer texts existed. The percentage of inscriptions with linguistic content out of the total number of inscriptions will again vary based on the type of the civilization. For civilizations like the IVC, they will be very low.

S.no	Type of Civilization	Percentage of inscriptions with linguistic content out of the total number of inscriptions	Percentage of long inscriptions out of a total number of inscriptions
1	Multilingual Trade based civilizations	Very low	Very low
2	Trade based civilizations	Low	Low
3	Monarchies	High	High

The sign distribution in the Indus script, using the current corpus of inscription is presented below. We also follow it up with a comparison with other scripts:

Frequency	Number of signs
1000 or more	1
999-500	1
499-100	31
99-50	34
49-10	86
9-2	152

Only once	112

This is not very different from the sign distribution frequencies in Proto-Cuneiform. In Proto-Cuneiform, astonishingly, the number of rare signs is much, much higher, in spite of the fact that this script was certainly used by a smaller number of people and over a lesser time span. In this respect, the Indus script would appear to be one step above Proto-cuneiform, which is a Proto-writing system. The total number of signs in Proto-Cuneiform is estimated to be 1617. [5]

Frequency	Number of signs
More than 100	104
99-10	370
9-2	610
Only once	530

In Proto-Elamite, Jacob L. Dahl states that there were 1900 signs, a number far higher than the Indus script. Out of these 1050 were attested only once, 8 between 100 and two hundred times, 8 between 200 and 400 times, 1 between 400 and 500 times, 1 between 500 and 700 times, and one over 700 times.

Such sign frequencies are not markedly different from the Indus script. In fact, going by this metric, it would appear that the Indus script which was the youngest of the three was the most stable. From our analysis, the Indus script should have had more signs that Proto-Cuneiform or Proto-Elamite because it was used by much larger numbers of people and over a longer period. This wasn't the case. It had much less signs than either. The sign usage is less skewed than either, notwithstanding the fact that the texts are short. [6] No texts in Proto-Elamite or Proto-Cuneiform display sign repetition in the manner the Dholavira signboard does. This would imply that the Indus script was a step above Proto-Cuneiform and Proto-Elamite. This would also imply that the Indus script was used more befittingly.

S.no	Script	Age	Number of signs	Proof available that pairing and ordering of signs played an important role	Long inscriptions	Inscriptions with sign repetition	Rare signs
1	Proto-Cuneiform	3200 BC (oldest)	1617	Yes	Yes	No. only Mature Cuneiform was true writing	High
2	Proto-Elamite	2900 BC	1900	Yes	Yes	No. only Linear Elamite glyphs display sign repetition	High
3	Indus script	2600 BC (The youngest)	1348 extrapolated	Yes	No	Yes. Dholavira signboard	High but not as high as Proto-Cuneiform and Proto-Elamite

The number of signs is one determinant of the nature of the Indus script because it would imply that the Indus script was based on the other two, and developed after Proto-cuneiform and Proto-Elamite. Only the signs would have changed as a result of the process of ethno genesis. Thus, the Harappans would have taken a state of the art script from West Asia in around 2600 BC when the

---

[5] *The Origins of Writing as a Problem of Historical Epistemology, Peter Damerow Cuneiform Digital Library Journal 2006:1*

[6] *Proto-Elamite sign frequencies, Jacob L.Dahl, Cuneiform Digital Library Bulletin 2002:1 version 29 April 2002*

magical transformation from the early to the mature Indus phase began and made changes to suit their requirements. Some changes, as we explained, would have been deliberate. It would also have been used differently, but this does not impact the nature of the script at all. Thus, the Indus script would have been used befittingly and would have most certainly been used to inscribe longer texts. The Dholavira signboard implies again, that it was one step above Proto-cuneiform and Proto-Elamite as the other two do not consist of any inscriptions displaying sign repetition. This type of an intra-region approach can stand us in extremely good stead as data expands and time progresses: it can then be used to solve riddles further west as well.

This Indus script is closely related to these scripts and is obviously as vastly different from Vinca Symbols, Tartaria tablets and Pictish stones as chalk is from cheese. Theories linking the Indus script to such symbol systems are complete non-starters and we have discussed this in detail in our previous paper. This, when taken together with all the other factors discussed in both our papers proves that it was a log-syllabic script: occam's razor applies, as always. It would be very obvious that Proto-Elamite was related to the Indus script, and a sign-by-sign comparison was already done by Fairservis and others. It was obvious that the split up happened when the transition from proto-Elamite to Linear Elamite was taking place and after it had reached the linguistic stage. Kudurru stones, on the other hand, are much later, dating to between the 16th and 13th century BC and a derivation of the Indus script from Kudurru stones, or the other way round is well-nigh impossible. Kudurru stones contained contracts as well, and these were usually in Cuneiform. What logical links does one see between Kudurru stones and the Indus script? The Akkadian empire, likewise, began later than the IVC mature phase, and the Indus script was clearly developed in 2700-2600 BC. This would rule out any source other that Proto-linear Elamite as a source for the Indus script.

We will now develop a method to predict the existence of a lost corpus of texts in the Indus which in addition to the Dholavira signboard will provide irrefutable evidence for the existence of a lost corpus of texts as archeological data begins to expand.

First step: To predict the total number of signs in the Indus script using a Bell curve

The very first step will be to predict the Total number of signs in the Indus script using a bell curve.The total number of signs in the Indus script has been arrived at using a bell curve – 1 new sign per 25 new inscriptions found at the current rate– one of the higher if not the highest figures Farmer mentions. This works out to a total of 1348 signs. This would include genuine singletons.

We include genuine singletons in the number of signs arrived at. We however assume that genuine singletons also decrease at exactly the same rate as other signs with the increase in number of inscriptions discovered. This assumption will not be valid in reality as singletons may continue to crop up at a constant rate. This, will not however change the total number of signs arrived at; we have arrived at a total number of 1348 signs, a figure far less than Proto-Cuneiform.

Our observations are as follows:

(1) The number of signs does not vary significantly from Proto-Cuneiform or from Proto-Elamite. As a matter of fact, the total number of signs in Proto-Elamite and Proto-Cuneiform is observed to be far greater at 1617 & 1900 signs, although these were used by smaller numbers of people over a smaller area over a much lower timespan. The number of rare signs in the Indus script is less.
(2) The total number of signs will not increase by a significant degree even if a lost corpus existed
(3) However, the Total actual number of signs in the Indus script will be equal to
Total Number of signs arrived at this method – Genuine singletons i.e. 1348-S

Percentage of the Indus excavated	Number of new signs	Percentage of the Indus excavated	Number of new signs	Percentage of the Indus excavated	Number of new signs
0-5%	417	35-40%	54	70-75%	16
5-10%	160	40-45%	45	75-80%	13
10-15%	134	45-50%	38	80-85%	10
15-20%	110	50-55%	33	85-90%	7
20-25%	92	55-60%	29	90-95%	4
25-30%	76	60-65%	25	95-100%	2
30-35%	63	65-70%	20	Total Number of Signs	1348

Therefore, the number of adjusted signs in Proto-Elamite is 1617 + Xa + Ya

Therefore, the number of adjusted signs in Proto-Cuneiform is 1900 + Xb + Yb

Where X is the differential added for the size of the IVC and Y is the differential added for the time span of 7 centuries which denotes the time span of the usage of the script. While a reliable estimate cannot be made of X and Y, they would be substantial in the eyes of most. We may note that these figures are added to the two scripts to make it comparable to the Indus script which was used for a longer time span and over a wider area. In other words, if these scripts were used in the IVC, they would have had a larger number of signs because the number of bonafide singletons and defunct signs arising due to internal changes in the script over seven centuries and usage over a much, much larger region.

Let us assume a conservative differential of 150 signs: this gives us 1767 signs for proto-Elamite and 2050 signs for Proto-cuneiform.

But the figure of 1348 representing the total extrapolated signs in the Indus script is a far smaller number, (It would remain significantly smaller even if it is rounded off to 1500 or even 1600 signs) naturally implying along with other factors that the Indus script was ahead of these two in terms of evolution, and it was also younger than these. This again would imply that a lost corpus existed when considered along with several other factors which we have discussed as the Harappans would have taken the state-of-the-art script further west and modified it to suit local needs. Even the most conservative scholar, would, after studying all the evidence that has been present in this and our earlier paper, concede that the Indus script was a logo-syllabic script that had reached a mature stage of evolution, shorter texts notwithstanding, and the 'non-script' adherents, whatever that may mean, have now very little to make their case. This test will like all others fall fully flat for Vinca symbols and Pictish stones.

Second step: The Isolation of non-bonafide low frequency signs and singletons

In the second step, we shall attempt to isolate non-bonafide low frequency signs and singletons. We shall do this as follows.

- Low frequency signs showing up across a wide geographical region are not bona fide low frequency signs. (Very reliable test for low frequency signs).
- Singletons on mass-produced objects are not bonafide singletons because they would have had to be read by many people.

- Low frequency signs or singletons bearing a close resemblance to Proto-Elamite signs or to signs in other contemporaneous scripts will not be bona fide singletons or low frequency signs because this would imply that these had to be used commonly. (very reliable test for singletons)
- Low frequency signs appearing several times in different inscriptions always displaying conditional entropy are not bonafide low frequency signs (reliable test for low frequency signs)
- Low frequency signs or singletons which are variants of other signs will not be bona fide singletons or low frequency signs.(fairly reliable test for singletons)
- Signs such as anthropomorphic ones may not be bonafide singletons (fairly reliable test for singletons).
- Sings bearing close similarity to other signs i.e. those forming a part of a group. (a less reliable test for singletons)
- Signs appearing in successor cultures may not be bonafide singletons.(a fairly reliable test for singletons)
- Highly abstract singletons may not be bona fide singletons – this is not a reliable test but can be combined with other tests to determine how bonafide a singleton is. Singletons can be ranked on a scale of 0 to 5, 0 being pictorial and 5 being abstract. (We do not claim this is a reliable test – a lost corpus can be inferred due to other reasons)
- Low frequency signs appearing in identical glyphs which were not mass-produced from the same mould may be bonafide low frequency signs. This test has to be made in conjunction with other tests.
- Singletons appearing as a part of a compound sign may not be bonafide singletons.
- Singletons or low frequently sings always appearing alone may be bonafide singletons.
- The changes that a singleton appearing in a longer glyph is a bonafide singleton is lower than it may be if it is appearing in a shorter glyph.
- The probability that a singleton appearing in a glyph with a linear arrangement is a bonafide singleton is much lower than it may be if it is appearing in a glyph with a non-linear arrangement.
- A singleton appearing in the middle of a glyph may not be a bonafide singleton.
- A singleton appearing in a glyph with sign repetition may not be a bonafide singleton. This is a particularly strong test.
- Any other signs that can be dropped out of the bonafide list – to provide justification.
- Each singleton must be assessed against multiple parameters to make a final decision.
- Tests must get stricter and stricter as data increases; in the early days tests need be necessarily strict. In other words, the probability that a singleton will remain a singleton in the early days is extremely low.

These signs can be removed from the list of singletons or low frequency signs. A preliminary analysis by Bryan Wells and other scholars would show that a large number of low frequency signs and singletons are not low frequency signs and singletons at all. In order to do this, we may prepare a singleton assessment checklist which tests all singletons and low frequency signs against all attributes mentioned above. This would help us categorize singletons into bonafide and non-bonafide singletons. Bonafide singletons may also be classified on the basis of probability that they are bonafide singletons, the rest can be categorized as doubtful. For example, only signs analogous to the "duck in pond" sign may be high-probability singletons. Many others bonafide singletons may be doubtful bonafide singletons, but one may like to wait for hard evidence to manifest itself.

As assessment of singletons and low frequency signs can proceed as follows:

Name of assessor						
Date						
Assessment of singletons:						
S.no	Indus sign:	Inscription code:	Location of discovery	Attribute	Assessment against Attribute	Overall assessment
1						

2						
				Summary	Bonafide singletons	
					Non-bonafide Singletons	
					Doubtful Singletons	
Assessment of low frequency signs						
S.no	Indus sign:	List of Inscription codes:	Discovery locations	Attribute	Assessment against Attribute	Overall assessment
1						
2						
				Summary	Bonafide singletons	
					Non-bonafide Singletons	
					Doubtful Singletons	

A summary table would give the following results:

Assessor No	Name	Number of bonafide singletons	Number of doubtful singletons	Number of non-bonafide singletons	Assumptions and remarks
A					
B					
C					

Putting all these results in a summary, we get the following results:

Frequency	Sign distribution existing corpus	Sign distribution 10%	Sign distribution 15% -95%	Sign distribution extrapolated 100%
1000 or more	1			
999-500	1			
499-100	31			
99-50	34			
49-10	86			
9-2 (M)	152-X			

9-2 (non-bonafide low frequency signs) (N)	X			
Only once (P)	112-Y			
Only once (non-bonafide singletons) (Q)	Y			
Incremental new signs (X)	Nil	160		
Incremental low frequency signs and singletons which may not be singletons and low frequency signs (Y)	Nil			
Incremental low frequency signs and singletons which may not be singletons and low frequency signs/ Incremental new signs (Y/X)				
Doubtful low frequency signs plus singletons divided by total number of low frequency signs and singletons (N+Q)/(M+P)				
Total number of signs (This has been extrapolated using a bell curve – 1 new sign per 25 inscriptions – one of the higher if not the highest figures Farmer mentions)	417	577		1348

Third step: Arriving at predicted Sign Distribution Frequencies

Sign distribution frequencies for each interval can be arrived at very easily. This may be done as follows

- Identify the number of new singletons likely to be discovered at each interval. We have already done this in the first step.
- Split up existing low frequency signs into bona fide low frequency signs and others using the criteria that we have identified in the previous step.
- Split up existing singletons into bona fide singletons and others using the criteria that we have identified at the previous step.
- The frequency for bona fide singletons and low frequency signs will not increase as the corpus of inscriptions increases.
- Predict the total number of signs under each of the frequencies for each of the intervals. One may also study how the usage of low frequency signs and singletons which may not actually be singletons or low frequency signs due to the reasons discussed, increases as the number of inscriptions increases. Does the number of singletons or low frequency signs which are not bona fide increase linearly, or at a faster rate or at a slower rate as the number of inscriptions increase? This input will be crucial to make an accurate prediction of the number of signs expected at each frequency.
- The total must tie up with the number of inscriptions expected at an interval multiplied by the average number of signs or 4.6.

$$\sum_{1}^{x} [f(k) \times f] = \sum_{1}^{x} [n(i) \times I]$$

Where I is the average length of an inscription, n(i) = number of inscriptions, f = frequency, f(k) = number of signs in a particular frequency.
- Refine the values as data increases. This will give us an idea of how accurate our computations were.

If the number of singletons or low frequency signs which may not be low frequency signs or singletons increases or remains large (Signs which are singletons or low frequency signs but do not appear to be upon closer scrutiny (reasons discussed above) to be legitimate singletons or low-

frequency signs), then it would certainly imply that a large corpus of texts existed. It would be illuminating to analyze the increase of the number of such signs as our data expands. These signs can tell us a lot about the script. These are the little fellows one must always monitor.

Fig2. If the number of new non-bonafide singletons being discovered declines slowly or at a moderate pace, it means there is better chance of a lost corpus being discovered. The reliability of this test increases as the percentage of Indus sites discovered increases. Therefore the possibility of statistical error is inversely proportional to the percentage of Indus sites discovered. There may be sudden and rapid drops of non-bonafide singletons after the discovery of hypothetical corpuses say C1,C2,C3,C4,C5 etc, but only the pattern of decrease thereafter has to be evaluated and studied to determine the probability of discovery of further hypothetical corpuses.
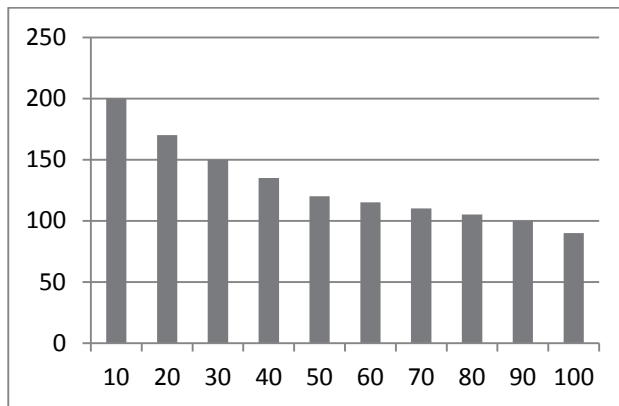

Fig.3 If the number of new non-bonafide singletons being discovered declines rapidly, it does not imply automatically there is little chance of a lost corpus existing However, this may well be a statistical error as a corpus can be inferred due to other reasons, and this is an extremely unlikely scenario. This scenario can also of course be used to show that a corpus on perishable materials existed or can be used to show that archaeological data is insufficient.
Either way, no one would now deny that a lost corpus existed. This reliability of this test increases as the percentage of the Indus sites discovered increases. Therefore the possibility of statistical error is inversely proportional to the percentage of Indus sites discovered. There may be sudden and rapid drops of non-bonafide singletons after the discovery of hypothetical corpuses say C1,C2,C3,C4,C5 etc, but only the pattern of decrease thereafter has to be evaluated and studied to determine the probability of discovery of further hypothetical corpuses.
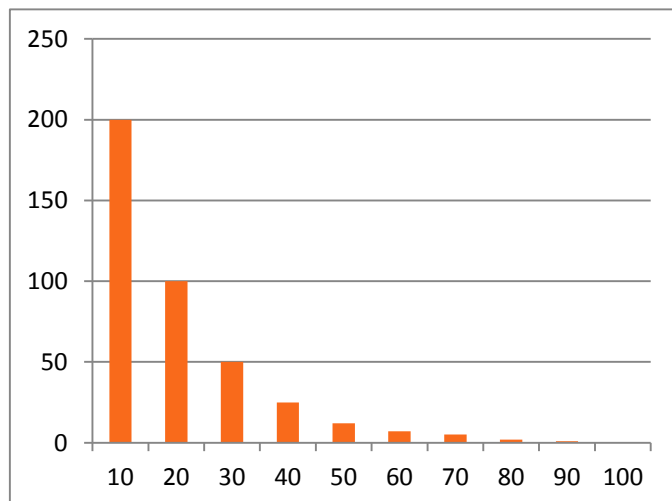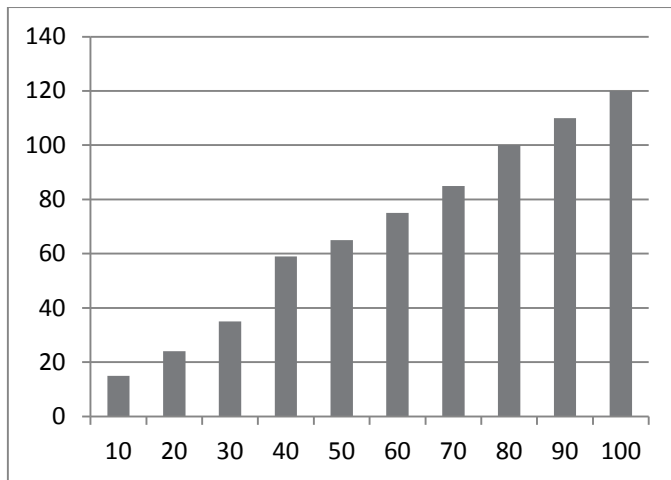

Fig 4. If bonafide singletons increase greatly as data expands, it would still imply that a lost corpus existed because of the other factors we have discussed. Remember to take contradictory evidence always into account. Bonafide singletons can be expected in a large civilization where the script was used by a large number of people. Moreover, the rare sign hypothesis is statistically unreliable at this stage. The percentage of rare signs does not vary significantly from other scripts as we have shown. Signs formerly identified as bonafide singletons may drop off the list. They may be marked as non-bonafide singletons if companions are discovered, but NEVER the other way around if the epistemology is correct.
The criteria for identifying singletons as was presented in our paper as bonafide is only indicative. if a lost corpus existed, many singletons could drop off the bonafide list. This could again imply that a hidden corpus existed, and that the classification of singletons as bonafide was wrong because evidence to make a reliable classification was inadequate.

Tell-tale evidence of a larger corpus

The existence of a much larger corpus of texts can be inferred in case any of the following conditions are observed:

- The number of singletons or low frequency signs which may not be genuine low frequency signs or singletons keeps on increasing.
- The number of singletons or low frequency signs which may not be genuine low frequency signs or singletons remains constant while the frequency of usage of more common signs keeps on increasing.
- The frequency of usage of such signs increases very slowly in comparison to the number of new inscriptions discovered.

The ratio that has to be monitored here is:

Total number of bonafide singletons - total number of singletons later discovered to be non-bonafide singletons i.e. signs dropping out of the list

Total number of singletons

OR

Total number of bonafide singletons - total number of singletons later discovered to be non-bonafide singletons i.e. signs dropping out of the list

Total number of signs

Examples of some other ratios that can be monitored in the context of increasing archaeological data are the ratios of doubtful singletons out of singletons classified as bonafide singletons.

This is in addition to the Dholavira signboard which itself implies that a lost corpus existed. This model will work much more reliably only as data expands but the advantage of this method is that it can be put to use early and the total number of sign under each frequency can be predicted. This can additionally be compared at every stage to the total number of signs of a logo-syllabic script of this type. Refer our earlier section for a more complete and meaningful discussion. There are three possible scenarios here:

Scenario A: The Indus script was not a stable writing system- this hypothesis was falsified in our earlier paper on multiple grounds and cannot even hold up to the mildest of scrutiny. Most other mainstream researchers have refuted this hypothesis already. Readers may wish to read the book by Michael Korvink 'The Indus script: A positional-statistical approach' which has approached the issue extremely systematically.

Scenario B: The Indus script was a proto-writing system.

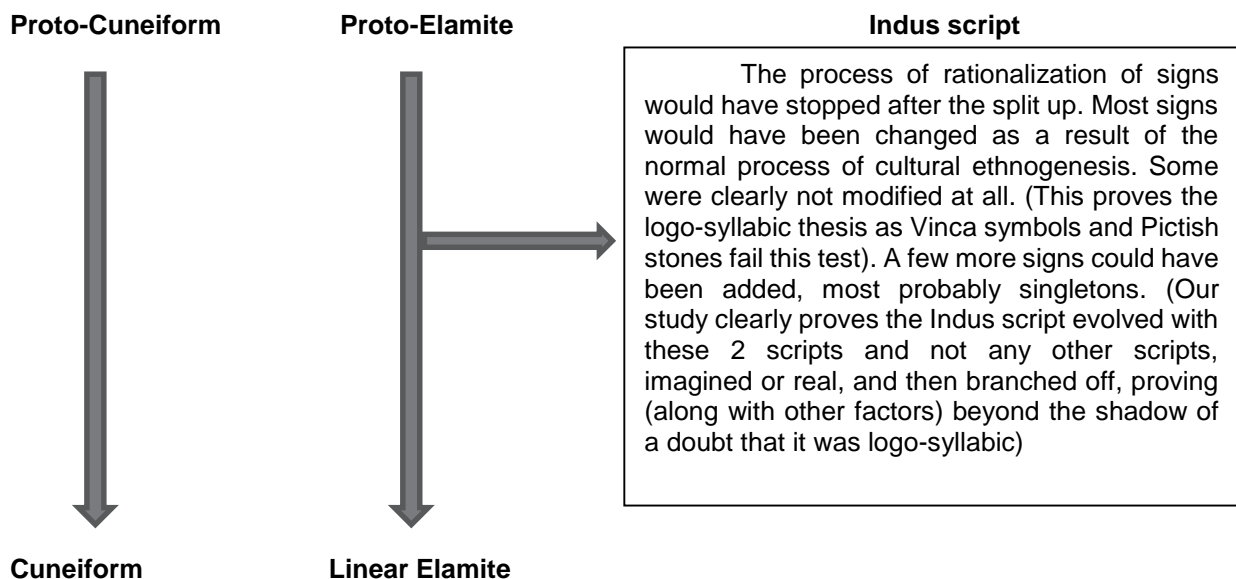Scenario C: Longer texts certainly existed in the Indus valley

Readers can make up their mind as to which of the three hypotheses is the most likely on the basis of our discussions in this and our earlier papers. The following sub-scenarios apply for scenario C.

Serial no	Sub- Scenario	Alternate hypothesis
1	Sub-scenario A: Lost corpus did exist	If no lost corpus is found, scenario B applies
2	Sub-Scenario B: Lost corpus existed on perishable materials (OR) archaeological data is insufficient	

'Once a singleton, always a singleton': How correct is this statement?

This statement is absurd because it assumes no further companions to singletons will be found, and as only 5% of the IVC has been excavated, this statement is practically meaningless. Signs formerly identified as bonafide singletons may drop off the list. They may be marked as non-bonafide singletons if companions are discovered, and this is a logical process of analysis of such singletons, but NEVER the other way around if the epistemology for identifying singletons as non–bonafide singletons is correct. If no more companions for non-bonafide singletons are discovered even after persistent digging, it would mean that archeological data is insufficient or a lost corpus of texts existed. (In addition, we have shown that the fear the Indus script had too many singletons is unfounded – refer this and our earlier paper.)

Sign redaction in scripts leading to a decline in number of signs in use

Proto-Cuneiform   **Proto-Elamite**   **Indus script**

> The process of rationalization of signs would have stopped after the split up. Most signs would have been changed as a result of the normal process of cultural ethnogenesis. Some were clearly not modified at all. (This proves the logo-syllabic thesis as Vinca symbols and Pictish stones fail this test). A few more signs could have been added, most probably singletons. (Our study clearly proves the Indus script evolved with these 2 scripts and not any other scripts, imagined or real, and then branched off, proving (along with other factors) beyond the shadow of a doubt that it was logo-syllabic)

Cuneiform   **Linear Elamite**

The origin of the Indus script

The idea that the Indus script derived from the rudimentary potters marks of the early Indus phase is untenable as we discussed extensively in our earlier paper on the Indus script, and cannot hold up to any scrutiny. The late Gregory Possehl says exactly this in his book, 'The Indus Civilization: A contemporary perspective'. We reproduce the relevant passages in the book..

"There is little, if any evidence for the beginning of writing in the early Harappan. Signs on pots, both pre- and post-firing, begin early, in stage two, but this is not writing, and some of it is probably simple potters marks, or marks of ownership. Logic suggests that the relatively developed writing system of Mature Harappan has its roots in the Early Harappan- Mature Harappan phase transition." [7]

Diverse usage of the Indus script

Most recent research demonstrates the Indus script to be versatile; there were at least four usages it was put to and this suggests different functions of the Indus script were used by different groups. All of them of course used exactly the same signs and combinations, and all of which most certainly required at least some amount of training for usage, but its usage to inscribe administrative records would have required more complex training and this need not have been imparted to other groups Neither Vinca symbols, Pictish stones or other 'symbol systems' display this kind of a versatility. These were:

S. no	Use	Used by	Required training?	Linguistic?	Was this the primary usage of the Indus script?
1	Personal identify	Elites	Yes. Both to read and write	Possibly not	Unlikely, if imported (This would imply that the elites used more complex communication as they were unlikely to have been trained only for this purpose)
2	Ritual use i.e. mass production of seals	Elites and shamans	Yes. Both to read and write	Possibly not	Unlikely, if imported
3	Trade and commerce	Traders and merchants	Yes. Both to read and write. Even a small class of Mesopotamian merchants would have been trained in the Indus script.	Certainly not if it had to be read by people who were outside the linguistic ambit of the Indus.	Yes
4	Administrative (The Dholavira signboard being the tip of the iceberg)	scribes	Required much greater training, and this would have been imparted to small groups	Yes. It's relationship to (various) spoken languages would have been a function of usage	Was one of the chief functions
5	Graffiti – only one such example was found and this	-		Maybe	No. this does however imply versatility.

[7] 'The Indus Civilization: A contemporary perspective'. Gregory Louis Possehl Altamira Press 2002.

	was short, but fairly large			

Evolution of Ideograms and Logograms to Logo-syllabic scripts

Instead of lumping scripts from different geographic zones and time periods into an unsavoury mess, it would be wise to analyze how ideograms morphed into logo-syllabic scripts; this was always happening in Old World Civilizations, with underlying motives and objectives being more or less same in all the three. However, the drivers for this could have operated differently in all three as the nature of civilization was different: the process would not have been normally abandoned once begun unless there was a major disruption such as the decline of the civilization due to other factors or the introduction of new technology such as alphabetic scripts. In our case, there are three scenarios:

Scenario	Description	Likelihood of the scenario	Implication
A	Was imported as a logo-syllabic script with predominantly non-linguistic usage but its full functionality employed by smaller groups. The implications of this scenario are: a. The Indus script was in all respects quite similar to the contemporaneous scripts of West Asia. b. It was only used differently because of the nature of the IVC. There is enough evidence to support this because the pairing and ordering of signs were similar to the scripts of West Asia. It had a similar number of signs as well. c. The Indus script played a critical role in the transformation of the early to the mature Harappan phase which happened rapidly because smaller groups of people who put the Indus script to different uses allowed this to happen.	Near certain	A lost corpus of texts certainly existed
B	Evolved in linguistic directions internally The implications of this are as follows: a. Evolved in linguistic directions internally without royal or centralized authority: not impossible but relatively less probable. Most Indus archeologists would even rule out this possibility entirely. In order for such an evolution to take place there has to be an evidence or royal authority or centralization, or at least priestly authority (this would hold good for the development of virtually all scripts) none of which were observed in the IVC (the IVC epitomized an anti-thesis of centralization with only some aspects such as the layout of the cities, some common cultural elements and of course the script being the same), automatically vindicating thesis A: this may account for Farmer's non-change thesis. The non-linguistic portion of the script which would have been used by much larger groups of people would have provided social cohesion in the intensely multi-lingual IVC (as Farmer rightly points out) and the script embodying full functionality	Low	A lost corpus would still have existed because the Dholavira signboard appears to represent a logo-syllabic script in mature form.(Even if thesis B is correct, there have to be several missing links and the context of its discovery proves the lost manuscript hypothesis)

	would have enabled its rapid script and would have facilitated record-keeping.		
	b. Does not account for the rapid transformation of the early to the mature Harappan phase.		
	c. Does not explain why the Harappans would leave out a vital feature when it was already available in contemporaneous scripts. We may bear in mind that they had taken other complex features such as the ordering and pairing of signs.		
	d. Readers may make up their mind as to which of these is more likely.		
	e. One can always argue that a few longer inscriptions occur towards the end of the IVC. But is this hypothesis statistically reliable at this stage with the low quantum of data? Have these been reliably been carbon dated? Does this hypothesis account for contradictory evidence? We leave it again to readers to judge.		
C	Had not yet reached the linguistic stage	Impossible because such a hypothesis would not even take all existing evidence into account	A lost corpus could not have existed.

What is the implication if the percentage of bonafide singletons out of the total number of singletons is high?

This ratio is statistically unreliable at this stage because of low quantum of archeological data. If would be useful to monitor this ratio over a period. Even if the ratio of bonafide singletons is to the total number of singletons remains stubbornly high as data increases, appeals and entreaties that a lost corpus of texts could not have existed will prove extremely chimerical: A lost corpus of can still be assumed due to other reasons which were discussed in the paper. Furthermore, bonafide singletons will drop off the list if a lost corpus did exist as our criteria for classifying signs as bonafide singletons was only indicative.

Three-tier model to research the Indus script

Future studies must split up the study of the Indus script into the following three categories instead of attempting to decipher the Indus script directly (a) A study of the nature of the script using a multi-disciplinary approach (b) A study of the usage of the script (c) A decipherment of the Indus script only if it is possible using a given corpus. Such analysis-driven approaches will stand us in very good stead always, and would be preferred to decipherment-based approaches.

Fig 5. This inscription is of the same length as the Dholavira signboard. However, the signs are non-repeating. This kind of dual usage clearly shows that the Indus script was logo-syllabic

The late Dr Gregory Possehl had echoed similar sentiments in 'The Indus Civilization: A contemporary perspective' [8]

"It may have been that the Indus scribes used a perishable material such as cloth or bark to compose longer documents. Both materials served as paper in historical times. There is always a chance that a large corpus of long inscriptions in the Indus script will emerge, perhaps on imperishable materials. Bilingual texts might also appear. But after a lot of digging, there is no evidence for them today, and this may suggest that we know the principal cultural contexts of Indus writing which is quite unlike Mesopotamia and Egypt. The context of early writing in these civilizations or archaic states was administration. In the Indus it seems to be personal identity. Professor J.Baines, an Egyptologist at Oxford University, made an intriguing observation pertaining to Indus texts written on perishable materials. There are many unique low frequency signs in the corpus of inscriptions on non-perishable materials and Baines noted that this may well be evidence that the inscribed materials recovered are sufficiently specialized and that the balance expression of corpus of signs is not present. If this is true, then it implies that at least some of the unique and low frequency signs that are recorded are simply a reflection of the deep bias on the corpus of texts and if we had the whole range of texts, e.g. on perishable materials, this statistical point would vanish. A real writing system with 113 signs or one where half the pictograms occur less frequently than five times is odd. The point that Baines makes is not a complete answer but I think it is important and should be kept in mind by the students of this ancient script."

Some more common questions addressed

We now attempt to address some more common questions about the Indus script. Answers to common questions people ask are presented below:

1. How important is the linear arrangement of symbols in determining whether a symbol system may be construed as a writing system or not?

A) Conditional entropy cannot distinguish between Logo-graphic scripts and Logo-syllabic scripts. It can at least be used to prove the stability of the system. All logographic scripts are writing systems. The question should be "Did the Indus script reach the syllabic stage or not?". There is enough evidence that it did. There is no such thing as a "symbol system". The classifications are logo-graphic and logo-syllabic.

The term "Symbol system" is highly superfluous for the following reasons:

Reason Number one: Its progenitors acknowledge it was a system. What then is the difference between a symbol system and a writing system? The onus is on them to provide a justification for this.

Reason Number two: A term to describe such a system already exists. There is no need to invent new terminologies unless existing ones do not satisfy a particular need.

Reason Number three: All logographic scripts morphed into logo-syllabic scripts at this time. In the case of the Indus script the process was under way, and there is enough evidence for this. Therefore, the Indus script was a logo-graphic script.

2. If the symbols are pictographic, is it important that they be abstract and not realistic?

---

[8] The Indus Civilization : A contemporary perspective, Gregory L.Possehl

A) Abstraction is a sign of the maturity of the system. Non-abstraction does not construe proof of the non-stability of the system. Indus signs were more abstract than Egyptian Hieroglyphs.

Abstraction does however, prove the following

1. Abstraction proves the stability of the system as was the case with Vinca symbols which were just proto-writing.
2. It proves its intended widespread usage as was the case with Vinca symbols (unlike Egyptian Hieroglyphs which may have been deliberately kept non-abstract to prevent wide usage)
3. Abstraction does prove that it was a system, as some sort of formal training would have become necessary.

3. How important is the length of the "texts" in deciding whether a symbol system is writing or not?

A) Length = function (usage). This length of Indus inscriptions was rightly questioned by Farmer but is a function of several different factors discussed in the paper. This includes the nature of the IVC. The existence of longer texts can be inferred due to reasons discussed in the paper. Again, there is no such thing as "symbol system". There are two categories of writing, proto-writing and writing.

4. In order to judge whether something is writing, how important is it that symbols repeat?

A) Symbol repetition = Function (usage). The Dholavira signboard shows sign repetition. This is proof that the Indus script was logo-syllabic when judged along with the context of its discovery. Most other glyphs of similar length do not demonstrate symbol repetition but this is an accepted feature of a logo-syllabic script and shows a kind of dual usage. A few other (short) glyphs have demonstrated sign repetition.

5. How important is the number of distinct texts to decide that a symbol system is writing?

A) Relevant only if it is weighed against the size of the corpus and the contexts of inscriptions. The presence of the same signs and sign combinations in a vast geographical region can however, speak volumes about its stability. This proves that the Indus script was a writing system. Please refer the tomes of research done on the Indus script by Western authors. There were many distinct inscriptions in the Indus as discussed in our earlier paper. This does imply that the Indus script was a writing system. Given the very low quantum of archeological data, the number of unique inscriptions will increase as data increases. Mass production of inscriptions which is a unique Indus feature has brought down the ratio of unique inscriptions out of the total inscriptions but proving it to be a logo-syllabic script using different methods and much more importantly arguing for a larger corpus renders this argument meaningless.

The difference between the Old logo-syllabic script thesis (Logo-syllabic thesis A) and the New logo-syllabic script thesis (Logo-syllabic thesis B): fundamental differences between the old and the new thesis

The wheel has now turned full circle and we are back where we originally began. There are however many difference between the Old Logo-syllabic script thesis and the New logo-syllabic script thesis and the differences can be summarized as follows:

Sno	New logo-syllabic script thesis (Logo-syllabic thesis B)	Old logo-syllabic script thesis (Logo-syllabic thesis A)
1	Shows using valid epistemology that the Indus	Assumes the Indus script to be logo-syllabic based on supposed prefixes and infixes in the Indus script. Many scholars have used a structural analysis to prove that

	script had certainly reached the syllabic stage.	the Indus script was Dravidian or Indo-European. This thesis has many limitations because seals would have been non-linguistic and would have been used in trade. There may have been some exceptions to this rule. (The term Indo-European or Indo-Aryan are too ambiguous in today's parlance – for a very detailed discussion refer to the paper 'Syncretism and Acculturation in Ancient India –A new Nine Phase Acculturation model explaining the process of transfer of power from the Harappans to the Indo-Aryans Part two.
2	Shows using valid epistemology that a longer corpus of inscriptions existed.	The lost manuscript hypothesis has had adherents for several decades. This hypothesis has however been by and large irrelevant to decipherment efforts.
3	Supports the idea that most seals were non-linguistic.	Assumes seals were linguistic
4	Supports the idea that the IVC was multi-linguistic.	Assumes that the IVC was mono-lingual
55	Supports the Idea that progress can only come from a structural analysis of the script. A decipherment with the existing corpus is unlikely.	Supports the idea that a decipherment of the script is possible.
66	Does not consider the decipherment of the Indus script to be of major relevance in Indology as a whole; multi-disciplinary and India-specific research strategies must be the starting point.	Supports the idea that the decipherment of the Indus script is central to further progress in Indology.

The difference between the Old logo-syllabic script thesis (Logo-syllabic thesis A) and the new logo-syllabic script thesis (Logo-syllabic thesis B): other differences between the old and the new thesis

We also do not blindly assume that a Rosetta stone will be found. A historical analysis of the contacts between the IVC and Mesopotamia & Egypt and a detailed study of the transformation of Harappan to post-Harappan India as was examined in our papers 'Syncretism and Acculturation in Ancient India: A new Nine Phase acculturation model explaining the process of transfer of power from the Harappans to the Indo-Aryans' Part one and two, which was published in the ICFAI Journal of History and Culture in January 2009 and January 2010 [9] can help here. We had also proposed methods to reconstruct the languages spoken in the IVC in these papers.

Also, we do not claim that the Indus script was wholly indigenous in origin: there is no nationalism in any of our papers. We only claim that evidence is adequate to establish that it was a logo-syllabic script. We reject the hypothesis that the Indus script was derived from the rudimentary potter's marks which were more ubiquitous in less developed societies found in the pre-mature Harappan phase. This scenario is highly unlikely because the internal synergies or authority to carry out the whole scale

---

[9] 'Syncretism and Acculturation in Ancient India: A new Nine Phase acculturation model explaining the process of transfer of power from the Harappans to the Indo-Aryans Part one and two: Sujay Rao Mandavilli, ICFAI Journal of History and Culture in January 2009 and January 2010

development of a logo-syllabic script would have been absent. This theory does not account for similarities with West Asian scripts either. These are additional but very, very important differences between the Logo-syllabic thesis A (or the old logo-syllabic thesis) and Logo-syllabic thesis B (the new Logo-syllabic thesis). All future researchers must make a difference between the two (as they test additional and new evidence against these two theses).

Transformation of Logograms to fully evolved Logo-syllabic scripts

We can map the transformation of Logograms to fully evolved Logo-syllabic scripts as follows, and this model will be fully consistent with experiences culled from excavations from all over the world:

Stage	Description	Does this feature provide proof that a script has reached the linguistic stage?	Has this feature been observed in the Indus script?	Has this feature been observed in Vinca symbols?
Early	Visual puns	Yes	By inference	No
Early	Abstraction	No	Yes. These were more abstract than Egyptian hieroglyphs. This is a proof of its intended widespread use unlike Egyptian hieroglyphs which were for a restricted audience.	Yes
Middle	Pairing and order of signs	No	Yes , was a key component of the Indus code	No
Late	Systematic use of determinatives	Yes	Yes (Dholavira)	No
Late	The use of the Rebus principle	Yes	Not attested yet	No
Late	Acrophony	Yes	Not attested yet	No

Factors increasing the probability of a lost corpus

The factors increasing the probability of a lost corpus are below

Sno	Factor	Implication
1	Test of interchangeability between the Dholavira signboard and contemporaneous scripts	Shows that the script had to be used more befittingly
2	The frivolous nature of the Dholavira signboard which means that longer texts existed	Do
3	The fact that speech encoding could be read by a large number of elites given the gigantic size of the signboard and its conspicuous position atop a citadel	Do
4	The fact that the glyph though short represents a mature form of a logo-syllabic script that was very obviously designed with longer texts in mind.	Do
5	The multilingual nature of the IVC	Increases the ratio of short:long inscriptions and therefore the probability of a lost corpus

6	The nature of use of seals: i.e. for commerce and the crucial understanding that it was a trade based civilization	Increases the ratio of short:long inscriptions and therefore the probability of a lost corpus.
7	Low percentage of the IVC discovered @5%, this is the lowest by far among all the major civilizations	Increases the probability of a lost corpus when taken together with other factors
8	Sophistication of the civilization and the availability of wherewithal in the period in question	Increases the possibility of a lost corpus
9	The evolution of the Indus script which shared features with the contemporaneous scripts of the Middle East (as opposed to Vinca symbols) Refer this and our earlier paper.	Increases the possibility of a lost corpus when considered along with all the other factors.

There are three possible scenarios as regards the Indus script and we have already completed a balanced and logical assessment of the script in our earlier paper.

Scenario	Description	Conclusion
Scenario A	The Indus script wasn't even a writing system	This hypothesis is no longer valid given that so much research has been done on the Indus script and its structure both before and after 2005, all of which have disproven this hypothesis already and the test of conditional entropy can at least be used to prove the stability of the system if studied along with several other factors.
Scenario B	The Indus script had reached the syllabic stage	In all likelihood, this must have happened as admitted by Farmer et al themselves. Therefore, the Indus script can be considered logo-syllabic. From the example presented in the paper, speech encoding does appear to have been one of the functions of the script and this function appears to have been built into the script and this implies that it was used commonly.
Scenario C	Longer manuscripts existed	Yes. The case for this has never been stronger than it has been now. (a) All the arguments for this are presented in the paper (b) So little of the Indus has been excavated (c) The low 1:x ratio only increases the probability that longer texts existed Royal inscriptions: The question of Royal inscriptions doesn't arise given the fact that it wasn't ruled by monarchs Literary records: Possible in theory, but such records as a result of royal patronage or tutelage can be ruled out. An assessment of how and why people wrote in contemporary civilizations and a study of the transformation of Harappan to Post-Harappan India can help here. Administrative records: Almost certainly existed. The logic and the reasoning are presented in this paper.

Therefore, the verdict would be obvious to the uninitiated and the layman alike, and those basing their conclusions on the canons of sound judgment and common sense. The Indus script issue isn't as easy as meets the eye, and a moderately sized corpus of administrative records certainly did exist. The changes in patterns of sign distributions as excavations progress can tell us extremely little about the size of the corpus for reasons that will be extremely obvious to anybody: only a dynamic assessment of the advancement of the Indus and the hierarchy and heterarchy of occupations and professions can help us here.

A find has been claimed very recently in Kabul, Afghanistan but may turn out to be spurious, if it is true, it may be the first in line in the discovery of longer texts. In the interim, we can reliably state
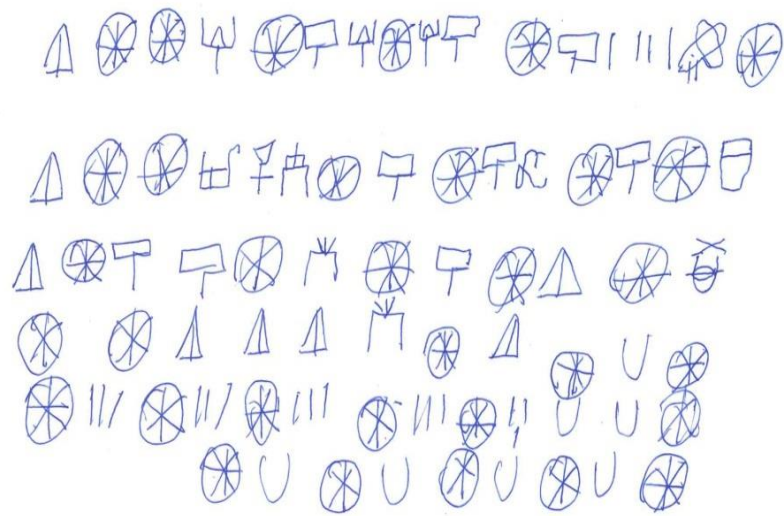
  (a) Existing evidence is sufficient to establish that it was true writing
  (b) Existing evidence is sufficient to establish that a lost corpus of texts existed

The direction of research: Indus archaeology as a driver of Indological research

Indus archeology must be indeed intensified because of the size of the IVC and its relative advancement. No other old world civilization known or undiscovered will match the size of the Indus valley civilization. The very fact that we now know that a lost corpus existed either on non-perishable or perishable materials must motivate archaeologists to excavate the IVC further, and Indus archaeology must become one of the pillars of Indological research. This can also improve our understanding of this Indus script and can also throw up interesting clues for the study of contemporaneous scripts further west. No one can deny that this will be of some importance to the rest of the world, and will even change the relative importance given to the IVC in history textbooks.

Someone must now pass the good news around. Imposing models used for the study of Egypt and Mesopotamia won't work: the sooner the Indologists realize this, the better. A Rosetta Stone may not surface for obvious reasons: readers may read our paper on the Aryan problem to figure out why: one approach might be to reconstruct the spoken languages of the Harappans as proposed in these papers; another may to compare the Indus script with the scripts found further west. Both may have to wait until a lost corpus emerges. In the interim, any Western or Indian scholar must insist on the following:

  (a) That the IVC is taught in all history textbooks around the world like its contemporaneous civilizations further West and be accorded its pride of place;
  (b) That the Indus script be discussed in the history of true writing systems;
  (c) That the Indus script and the near-certain existence of a lost corpus be discussed in all articles pertaining to the IVC, general literature and in textbooks.as well;
  (d) That the IVC is not omitted in any encyclopedic literature or books discussing human civilizations; those omitting it will be committing a cardinal error.
  (e) That the fact that the Indus script is now known to be true writing become a driver in archeological research as a lost corpus will certainly be found some day if inscribed on non-perishable materials. It no such corpus is found, it would mean that longer texts were inscribed on perishable materials. We have included a detailed discussion on writing materials in our previous paper. This alone must motivate more researchers to take up research on ancient India apart from other factors such as those discussed in our earlier paper which presented a comprehensive solution to the Aryan problem.

Fig 6. What a longer inscription (these most certainly would have existed to state the very,very obvious) may have looked like: based on an extrapolation of signs from the Dholavira signboard which shows a random outwardly repetition as could be expected in any true logo-syllabic script: what other proof could one want that a lost corpus existed?

Whatever, one might say about Pictish stones or other "symbol systems", the Indus script is clearly not a part of this category. The term 'symbols' may not have much weightage because there is a perfectly legitimate term for them: ideograms, pictograms and logograms. If Farmer wants to use the term 'symbol' to represent ideograms which do not form a system at all, he may be but only partly right. The Vinca symbols do form some kind of a system even though 85% of its symbols appear in isolation unlike the Indus script. The Vinca symbol system does display abstraction and standardization which implies that training was required to read the symbols and that it was designed to be used over a wide geographical area. This term again would not hold any good for the Indus script if it had reached the linguistic stage. Refer our earlier paper, 'The reconfirmation and reinforcement of the Indus script thesis' which was published in the ICFAI journal of History and Culture in January 2011. The theory that the Indus script was not even a writing system has failed to stand all tests carried out since 2004 by Western scholars of eminent repute and was a complete non-starter. This is apart from its blatant political overtones, and the highly undesirable, irresponsible & unhealthy penchant of bringing an extreme amount of personal prejudice & bias into science. We have furthermore, reproduced quotes from scholars of UNIMPEACHABLE INTEGRITY such as Jane McIntosh and the late Gregory Possehl in support of our stand. The compartmentalization of writing systems into symbol systems and scripts is

unhealthy and practically meaningless for many reasons, and future researchers would be well-advised to avoid the term "symbol system" as far as possible due to the following fundamental reasons:

(a) Perfectly legitimate terms to describe symbols exist i.e. logograms, pictograms and ideograms exist.

(b) Pictograms were always evolving into logo-syllabic scripts at the time the Indus valley civilization flourished and in cases like the Vinca symbols this transition could not have certainly happened because these were much, much older and pertained to a period when true writing did not exist.

(c) Compartmentalizing writing systems into scripts and symbol systems pre-empts further meaningful research as a multi-disciplinary approach must be applied for each of the scripts in question to address the issue of whether scripts were logosyllabic or not. The great Russian scholar Yuri Knorozov, who convincingly deciphered Mayan hieroglyphs, likewise argued that scripts were rarely non-linguistic. He may have been partially right because speech-encoding would have been an important feature, and if they had access to it, they would have borrowed it.

(d) It automatically introduces bias and prejudice. In a narrow field of study where the number of players is small, and ideas are generally not challenged easily or quickly, scholars must be more responsible and accountable for their actions. One can use terms such as symbol systems and scripts recklessly to mislead the general public. One can compartmentalize civilizations based on his or her dislikes: as we observed, the Indus had a unique interpretation of literacy, such as the mass production of writing and the ability of large sections of the Indus population to read the Indus script. These kinds of approaches, and the introduction of a high degree of personal bias sets a bad precedent not just for Indology, but for science as a whole and are very unhealthy not only for the simple reason that they mislead the general public but also because they preclude multidisciplinary approaches from being adopted and affect research in writing systems in general.

(e) Given the fact that we now know it was a logo-syllabic script, this model can be used to research scripts further west as data expands.

Therefore, the question should be 'Did the Indus script reach the linguistic stage or not'. There is enough evidence to adduce that it did.

As so little of the Indus has been excavated, drawing a priori conclusions about "non-literacy" (whatever that may mean) is absurd, because it is against the principles of reasoning and epistemology & is also a blatant misuse of the term besides containing leaps of logic and non-sequitors. The IVC was literate in full conformity with well-accepted definitions of literacy. We have also shown in our earlier paper that none of Farmer's conclusions on the Indus script were correct, length notwithstanding: but that can prove nothing; one can conclude that modern Indian scripts are not scripts on the basis of the length of 'inscriptions' on signboards and shop signs because there were the only 'inscriptions' on non-perishable materials and these were short in length; besides the brevity of the inscriptions is known already. We have also validly and logically shown that a lost corpus of texts existed in the Indus even with existing evidence and that existing evidence is adequate to establish that is was true writing. This does not certainly violate any principles of epistemology. All these, were taken and studied together, must be a great moment of epiphany for us and help us comfortably sidestep all perverse shenanigans; More evidence must, of course, be obtained, (as unlike in contemporaneous civilizations, Indus archaeology is in its infancy in any case, and therefore these approaches stand is much better stead) and Indus archeology, needless to say must become one of the drivers of 21st century Indology.
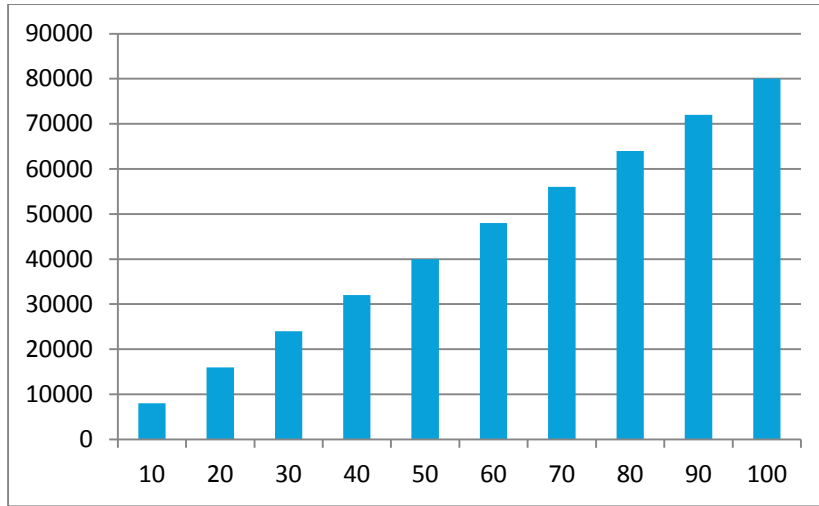
Fig 7. Non-linguistic seals in all likelihood will keep showing up at a constant rate. The X axis represents the percentage of the Indus discovered while the Y axis depicts the number of inscriptions discovered (cumulative).
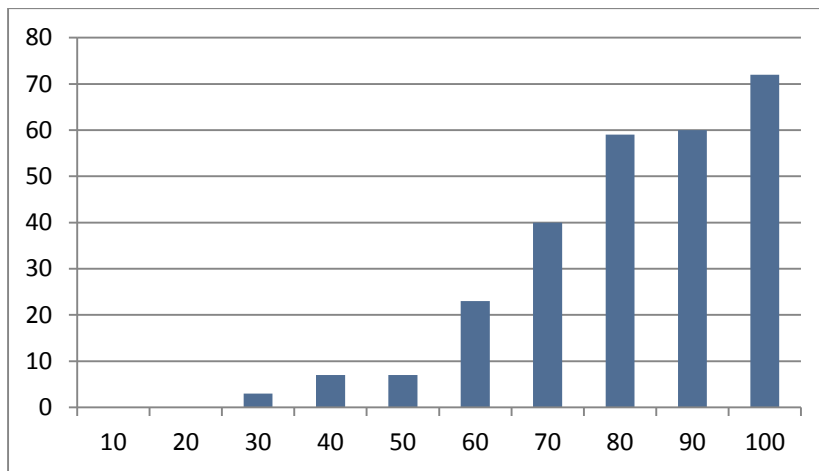

Fig 8. Surviving non-linguistic texts may not throw up at a constant rate because these were perhaps few and far in between and also were found closer together. All were probably not as closely tied to speech as the Dholavira signboard but would have made use of features such as those discussed in the paper which means they were logo-syllabic. We say this because the relationship with spoken language would have been a function of usage, and the tendency would have been for inscriptions to have used as many non-linguistic signs as possible so that these could have been ready by a larger number of people. The Dholavira signboard still suggests that speech-encoding was a primary function of the script, and this implies that such examples also would have existed. The length of the text can alone mean nothing. Rongorongo inscriptions (which were long) were probably not closely tied to speech either. We can also therefore argue that it is not necessary that the signs used in longer inscriptions were the most frequently used ones. This does not upset the principles of this paper; the discovery of non-bonafide singletons can still be used to predict the discovery of a lost corpus.

The end result of the decade-long debate, apart from the fact that it has phlebotomized the issue, put an end to pseudo-decipherments that quite literally marred the field in earlier decades, and that it has allowed the logo-syllabic thesis to be modified, is that it has allowed us to demonstrate that even existing evidence is adequate to prove that a lost corpus of texts existed in the Indus, and that this automatically boosts its pedigree, necessitates the expansion of archaeological research (because this will also solve riddles further West like helping us to understand the relationship between Proto-Elamite and Linear Elamite), apart from increasing the importance of old world civilizations in general. The theory that the Indus script was not a logo-syllabic script is, of course, defunct. And it deserves to be buried & forgotten, because the chances that it can be resurrected are virtually non-existent. Does anybody still beg to differ?