Breaking down the bullion. The compliance of bullion-currencies with official weight-systems in a case-study from the ancient Near East

Nicola Ialongo a, *, Agnese Vacca b, Luca Peyronel c

a Georg August Universität Göttingen, Seminar für Ur- und Frühgeschichte, Nikolausberger Weg 15, D-37073 Göttingen, Germany
b Università di Roma “La Sapienza”, Dipartimento di Scienze dell’Antichità, Via dei Volsci 122, 00185 Roma, Italy
c Libera Università di Lingue e Comunicazione, Dipartimento di Studi Umanistici, Via Carlo Bo 1, 20143 Milano, Italy

1. Introduction

The main question of this study concerns the means through which a bullion-currency can be effectively identified, and distinguished from other kinds of non-currency goods. In particular, we seek to verify whether a quantifiable relationship exists between bullion-currencies and standard weight-systems. The hypothesis is that the weight values of bullion — both complete and fragmented — should comply with standard weight-systems. Therefore, statistical tests on sets of bullion and sets of balance weights, from the same economic network, are expected to give comparable results.

The research focuses on a Middle Bronze Age hacksilver hoard from the city of Ebla, in Inner Syria (Fig. 1). Several hoards of silver bullion, dated to the Bronze Age, have been found in Near Eastern sites. Unfortunately, the weight values of silver pieces are generally absent, and the graphic documentation is often lacking. Silver hoards occur in public and private contexts, with varying in- terpretations as deposits of value or ritual depositions (Peyronel, 2010: 928, and fn 12–13). Through the analysis of the silver hoard of Ebla we aim at reaching a deeper understanding of the modes through which scrap silver and silver objects effectively circulated, before being hoarded.

The Ebla hoard contains 171 silver items, mainly fragmented pieces and ingots. The items included in the hoard, with the only possible exception of a silver bead, are neither utilitarian objects nor ornaments of any kind. Considering that texts often mention payments being made in weighed silver (Paoletti, 2008; Pomponio, 2003; Milano, 2003; Arkhipov, 2012: 12), it can be assumed that whatever the purpose of the deposition — the state in which the objects were recovered was likely the same state in which they actually circulated: ingots and scraps, passing from hand to hand as a means of payment, eventually collected and buried underground. Cuneiform texts from Mari mention different metal items circulating in the form of disks (kakkar[t]um), sheets (k’l’um), or lumps (kubulum), and specific terms might also refer to silver and gold scraps (sibirtum), as well as to portions (sankuttum) of metals (Arkhipov, 2012: 17–21). All these terms find close correspondences with the items included in the Ebla hoard. In this perspective, the Ebla hoard represents a sample of circulating hacksilver, a form of bullion-currency whose exchange value had to be assessed through weighing, and formally acknowledged against standard frames of reference.

* Corresponding author.
E-mail addresses: nicola.ialongo@uni-goettingen.de (N. Ialongo), agnese.vacca@ gmail.com (A. Vacca), luca.peyronel@iulm.it (L. Peyronel).

https://doi.org/10.1016/j.jas.2018.01.002
0305-4403/© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
A comparative analytical framework is defined, to test the statistical properties of hacksilver against those of balance weights (see supplementary material). The Ebla hoard will be compared to five groups of balance weights, all dating to the Middle Bronze Age (hereafter MBA or MB; c. 2000–1600 BC), from different sites: the Assyrian karum ("trade post") of Kültepe, in Anatolia; the city of Ebla, in Northern Syria; and the cities of Nippur, Larsa and Ur, in Southern Mesopotamia (Fig. 2). The study aims to be a contribution towards the definition of a general model for the identification of bullion- and commodity-currencies, that can be further developed to include different pre- and protohistoric economic systems.

2. Research questions

How far do official standards concur in determining the weight of bullion-currencies, and ultimately in producing the materiality we observe in the archaeological record?

Is there any regular pattern in bullion-currency samples that we can use to infer normatively-induced behaviour?

In practice, is the distribution of the weight values of bullion-currencies similar to those of balance weights in any significant way?

The answers to these questions are key to understand the process leading to the formation of the archaeological record related to any kind of bullion-currency. The results will help clarify the way...
bullion-currencies are shaped through economic transactions, and whether or not bullion, even when fragmented, is to be expected to comply with any existing standard weight-system. This problem is an important matter of debate, for example, in the metrology of the European Bronze Age, where copper is often thought to have been employed as a form of currency (Ialongo et al., 2015; Lenerz de Wilde, 1995; Malmer, 1992; Pare, 1999; Peroni, 1966, 1998; Primas, 1997; Sommerfeld, 1994; Sperber, 1993; Wiegel, 1994). Due to the generalized scarcity of fully-acknowledged balance weights in Europe (at least, outside of Greece: Cardarelli et al., 1997, 2001, 2004; Feth, 2014; Pare, 1999; Vilaça, 2003, 2013), in fact, several studies base their analyses on the assumption that copper fragments, as a form of bullion-currency, should conform to units and multiples of existing weight standards; official standards, therefore, are often inferred through the analysis of the distribution of weight values of metal objects, either complete or fragmented. But is this assumption valid in the first place?

The hacksilver hoard from Ebla was chosen to address the relationship between normative systems and bullion-currencies, in a relatively advanced economy: hacksilver, as a form of weight-based currency, can be expected to lie somewhere in between the theoretical exactness of balance weights and the indeterminacy of normal trade-goods. It can be expected to have been modelled and fragmented in order to approximate to exact multiples of a given unit, but, at the same time, nothing prevents that hacksilver was simply weighed and assigned a value a posteriori.

3. Theoretical framework

Balance weights and weight-regulated objects are the most direct material correlates of standard weight-systems. Two main factors concur in shaping the relationship between the two object-categories: first, a normative constraint, i.e., the need to comply with official standards; and second, the degree of approximation that is allowed to economic agents, before the normative constraints are violated.

The former is a necessary assumption in every attempt to reconstruct ancient systems of measurement. Exact units, however, are merely theoretical concepts, and theoretical exactness inevitably clashes with the inherent inaccuracy of real-world practice: for as accurate a measuring tool can be, it will always be affected by some degree of inaccuracy (Arkhipov, 2012: 183; Durand, 1987; Ialongo and Vanzetti, 2016; Joannes, 1989: 127). This problem is inherent to any measuring society, not only to ancient ones. Still, trade and exchange were never much affected by it: inaccuracy, in fact, is easily worked around through bargain, trust, convenience or even fraud (Ialongo et al., 2018). In practice, all these factors concur in increasing the statistical dispersion of archaeological samples: the theoretical exactness of balance weights and the indeterminacy of normal trade-goods. It can be expected to have been modelled and fragmented in order to approximate to exact multiples of a given unit, but, at the same time, nothing prevents that hacksilver was simply weighed and assigned a value a posteriori.

4. Materials and methods

4.1. Materials

The Ebla hoard1 was found beneath the floor (L3702) of a Middle Bronze II (c. 1800–1600 BC) house, located along the southern slope of the Acropolis, and probably part of a neighbourhood of craftsmen attached to the Royal Palace (Nigro, 2003: 347). The hoard was originally interpreted as a funerary assemblage of an under-floor adult burial (Baffi 1987: 4, Fig. 2: 6–11). While the archaeological context represents per se an exception (Peyronel, 2010: 931), the content of the jar is common to other Bronze Age hoards from the Ancient Near East, usually consisting of scrap, ingots and diverse metal objects. The silver was kept in an ovoid jar, filled with 171 silver items weighing 5043.5 g in total, roughly corresponding to ten Mesopotamian minas of c. 500 g (Fig. 3: supplementary material). The silver bullion includes 117 scraps, 53 ingots and 1 large bi-conical bead (Table 1). Small silver fragments are the most represented category (83 items), followed by complete and fragmentary, flat discoid ingots (47 specimens) and complete and fragmentary bar-ingot fragments (20 items) and rods (11 items). Thus, the inventory mostly includes different types of ingots (either complete or deliberately fragmented), while no jewels, or objects to be recycled are attested, other than a single silver bead and 5 folded sheets and 4 melted objects.

Five groups of balance weights were considered in the comparative analysis. All the samples are quantitatively significant, and pertain to important Bronze Age sites, covering a large territory ranging from Anatolia to Southern Mesopotamia (Fig. 2).

A total amount of 197 balance weights dating to MB I-II was collected at Ebla, 121 of which from stratified contexts (including the palace, temples, private houses and defensive structures), while the remaining 76 are from secondary contexts (Ascalone and Peyronel, 2006: 209–211, tab. 6.2). Only the complete exemplars were considered in the present analysis, for a total of 94 wt: 2 dating to MB I (c. 2000–1800 BC), 10 to MB II, and 82 coming from destruction layers of the MB II city (c. 1800–1600 BC) (Ascalone and Peyronel, 2006: 542–585).

The 162 balance weights from the Anatolian site of Kültepe were all retrieved from levels I-a-b and II of the Assyrian kuru, and date to MB I (Özgüç, 1986: 77–81; Dercksen, 1996: 80–89, 251–253; Ascalone and Peyronel, 2006: 410–420; Kulakoglu, 2017). The majority of balance weights is from domestic and funerary contexts of Assyrian merchants (in burials, they often occur in scale-sets together with scale-pan), as well as from metallurgical workshops. Three Southern Mesopotamian sites have been taken into account in the present analysis: Ur, Nippur and Larsa.

Balance weights from Ur (fully published in Hafford, 2012) account to 327 unbroken exemplars, of which 299 come from traceable contexts, mostly between the Early Bronze and the Late Bronze Age (2500–1200 BC). Nearly half of the sample (43.8%) was retrieved in funerary assemblages, while the remaining 56.2% pertains to domestic and public contexts, or has a generic provenance (Hafford, 2012: tab. 8). Interestingly, at least a few of these objects are possibly from the Diqdiqqeh, the craft and commercial quarter of Ur, located outside the city, and dated between the Early and the Middle Bronze Age (Hafford, 2012: 47; Ascalone and Peyronel, 2006: 445). The majority of balance weights from funerary contexts can be ascribed to the EBA and MBA periods (12 and 14 burials respectively; Hafford, 2012: table 8.40). In particular, MBA graves were excavated beneath the floor of private dwellings

---

1 The hoard was studied by Luca Peyronel in 2008, in the National Museum of Idlib, Syria.
dated to the Isin-Larsa period (c.a 2000–1800 BC) and yielded — as in the case of Kültepe — copper scale pans together with sets of hematite weights (Peyronel, 2000, 2011).

261 weights are recorded from the site of Nippur, of which 132 complete exemplars have been included in the analysis (Hafford, 2005: Table 3). The archaeological context is known for only 64 weights, dating to the late EBA (Ur III) and MBA (Isin-Larsa and Old Babylonian) periods (c. 2100–1600 BC); they are all from private dwellings (Hafford, 2005: 364–366).

Finally, 67 balance weights from Larsa, dating to MB I, were retrieved inside the so-called "goldsmith’s hoard", buried underneath the most recent floor of a room (Court I, room 13) of the Ebabbar Temple (Arnaud et al., 1979). Other than the weights, the jar contained several jewels, precious stone beads, bronze tools, and silver scraps, together with administrative/economic tools including a cylinder seal, 18 cretulae with short inscriptions.
recording amounts (of silver?) in shekels, and one cuneiform table, providing the recap of the weighing operations (Ascalone and Peyronel, 2006: 451–464; Peyronel, 2010: 932). The cunea, sealed with the cylinder seal of Sin-uselli (an official of the weights bureau of Ur), and the small table are particularly interesting, since they attest the practice of official weighing carried out by the central administration (Arnaud et al., 1979: 17–18). Unfortunately, the weights of silver scrap are not reported in the publication, making it impossible to test the statistical properties of hacksilver against those of balance weights, and to assess the correspondence between the figures provided by the tablet and the total amount of bullion.

### 4.2. Methodological premise

The most common analytical techniques employed in the research on early weight-systems and weight-regulated currencies are the Frequency Distribution Analysis (FDA) and the Cosine Quantogram Analysis (CQA; Kendall, 1974). The former is an intuitive method, mainly employed in metrological studies on Bronze Age Europe (Cardarelli et al., 2004; Lenerz-de-Wilde, 1995; Melmer, 1992; Paris, 1999; Peroni, 1966, 1998; Sommerfeld, 1994; Sperber, 1993; Wiegel, 1994); the latter is more sophisticated, and has become a standard method in ancient weight-metrology (Bobokhyan, 2006; Hafford, 2005, 2012; Lo Schiavo, 2009; Pakkanen, 2011; Paris, 1999; Petruso, 1992; Pulak, 2000; Rahmstorf, 2010). While the effectiveness of CQA has been thoroughly tested in different research contexts, FDA is more problematic. The assumption about FDA is generally posed in the following terms: significant clusters in the distribution of weight values should correspond to integer multiples of a given unit of measurement; hence, it should be possible to infer the weight-system through the analysis of the mean values of clusters. However, FDA is a descriptive method, and does not allow to make inferences about statistical populations. FDA must be coupled with CQA (and vice-versa) in order to assess the variability of the sample (Kendall, 1974; Pakkanen, 2011). Furthermore, significant concentrations in the distributions of weight values are not necessarily related to the compliance with a single normative system, since balance-weights belonging to different systems can share the same clusters (Ialongo et al., 2018); they can depend, instead, on other forms of economic behaviour that, in most cases, have the purpose of overcoming the constraints imposed by the conversion between different normative systems, and to establish cross-system standards that are acknowledged through customary economic-relationships, rather than normatively enforced. The methodology of this study was designed to address such a duality, and the results of the analyses will be employed to trace a general framework that accounts for different interpretive aspects.

### Table 1

<table>
<thead>
<tr>
<th>Alleged units</th>
<th>Agean</th>
<th>Syrian</th>
<th>Mesopotamian</th>
<th>Levantine</th>
</tr>
</thead>
<tbody>
<tr>
<td>reference to Fig. 3</td>
<td>#1</td>
<td>#2</td>
<td>#3</td>
<td>#4</td>
</tr>
<tr>
<td>Ebla Hacksilver</td>
<td>6.90</td>
<td>7.60</td>
<td>8.60</td>
<td>9.50</td>
</tr>
<tr>
<td>Ebla Weights</td>
<td>6.80</td>
<td>7.40</td>
<td>8.30</td>
<td>9.10</td>
</tr>
<tr>
<td>Kültepe</td>
<td>6.90</td>
<td>7.50</td>
<td>8.20</td>
<td>9.00</td>
</tr>
<tr>
<td>Nippur</td>
<td>7.90</td>
<td>8.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larsa</td>
<td>6.90</td>
<td>8.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ur</td>
<td>7.70</td>
<td>8.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>6.88</td>
<td>7.62</td>
<td>8.33</td>
<td>9.20</td>
</tr>
<tr>
<td>CV</td>
<td>0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Our methodological framework is based on a simplification of the problem, and attempts to answer three specific questions.

1) Is it possible to recognize standard units in the distributions of hacksilver weight values?
2) How similar is the “quantal pattern” of hacksilver to those of balance weights?
3) Are there significant clusters of values that can be compared across different distributions of hacksilver and balance weights?

The questions are addressed through three different sets of analyses, respectively Cosine Quantogram Analysis, Correlation Analysis of quantograms and Frequency Distribution Analysis.

Finally, a note about alleged normative units. The debate on the exact identification of standard units in the Near East and in the Mediterranean is still ongoing, and some units are less widely accepted than others (Alberti and Parise, 2005; Ascalone and Peyronel, 2006: 17–49; Zaccagnini, 1999–2001, 2006). However, this study is not aimed at the reconstruction of theoretically-exact systems of measurement, but rather at comparing hacksilver and balance weights in order to look for common patterns. Therefore, the methodological framework was specifically designed to extract meaningful information from the samples, regardless of whether the individual objects belong to this or that theoretical system. When discussing our results, we provide an interpretive model that takes into account the simultaneous presence of different systems in the same contexts.

### 4.3. Cosine Quantogram Analysis

The first question is addressed through CQA, the so-called “Kendall Formula”. Based on a Fourier transform, CQA was devised by D.G. Kendall (1974) to test whether an observed measurement $X$ (in our case, a balance weight or a hacksilver fragment) is an integer multiple of a “quantum” $q$ plus a small error component $\varepsilon$. Since the method was designed to address the natural dispersion of real measurements, the error component $\varepsilon$ is a fundamental part of the analysis. In order to assess the accuracy of a quantum $q$ in producing a measurement $X$, $X$ is divided for $q$ and the remainder is tested. Positive results occur when $\varepsilon$ is close to either 0 or $q$, i.e., when $q$ is (close to) an integer fraction of $X$. The following formula calculates the score $\phi(q)$ for the tested error $\varepsilon$ of a given quantum $q$, where $N$ is the number of measurements in the sample:

$$\phi(q) = \sqrt{2/N} \sum_{i=1}^{n} \cos \left( \frac{2\pi i \varepsilon}{q} \right)$$

The value of $\phi(q)$ for each quantum accounts for the whole sample, through the summation of the scores of the error component $\varepsilon$. The result of the analysis is a list of values of $\phi(q)$ for every tested quantum; plotted against the quanta in ascending order, the values of $\phi(q)$ render a wavy line with positive and negative peaks (“quantogram”), where positive values identify probable quanta (hence, possible units) for the sample under analysis.

### 4.4. Correlation analysis of quantograms

While CQA is now established as a reliable tool for the identification of possible unit-values in uniform contexts, the possibility to use it to address the potential similarity between different datasets was never explored. CQA retrieves a positive or negative score for each tested quantum, producing the characteristic graph. It has been demonstrated that both positive and negative values are equally significant (Kendall, 1974); therefore, when exploring the similarity between the quantal patterns of different samples, both
Fig. 4. Quantograms of the analyzed samples. The vertical bands highlight the probable intervals of known shekels, for quanta with relatively high values of $\phi(q)$. The values breakdown is given in Table 1.
positive and negative values should be considered.

The Pearson's product-moment correlation coefficient ($r$) was chosen as a measure of similarity between two quantograms. However, the fuzziness that usually characterizes ancient metrological samples can be a source of bias, and must be dealt with. Therefore, the possibility to compare different quantograms was tested through Monte Carlo techniques, with the aim of assessing the likelihood that more or less high values for $r$ may result from chance, when correlating randomly generated variables (see Appendix 1). The tests show that the analysis tends to be significant for values of $r > 0.355$.

4.5. Frequency Distribution Analysis

Finally, we analyse the complete dataset in order to locate significant concentrations of weight values that can be compared across different samples of hacksilver and balance weights. The data are classified into binned distributions, and significant concentrations are identified. In practice, the analysis seeks to identify a series of peaks that are supposed to result from a random error chance around a predetermined value.

5. Analysis

5.1. Cosine Quantogram Analysis

For all samples, CQA is executed for 100 quanta ranging from 1 to 10, with an increment of 0.1 (i.e., the range within which the known standard units and their submultiples are expected to be located). The hacksilver sample is considered entirely, while only completely preserved balance weights are included in the analyses. Finally, no measurements below 1 g are considered, because of the bias deriving from either modern measurement and/or ancient inaccuracy.

In all quantograms, the relevant peaks are distributed throughout the graph (Fig. 4). Four significant value-ranges can be identified between c.a. 6 g and 10 g, each one easily comparable with one of the four basic units (shekel) that are commonly accepted for the Bronze Age Near East (Table 1).

The alleged “Mesopotamian” unit of 8.4 g (as well as its half) is very well represented in all samples, except that from Larsa. The divergence of Larsa from the other samples is probably due to the unusually high occurrence of the value 2/3 (i.e., 5.6 g), altering the results of the CQA. This interpretation is supported by the high peak around c.a. 2.8 g (1/3 of the unit), which is highly represented and shared by all samples. The “Syrian” unit of 7.8 g is also widely

<table>
<thead>
<tr>
<th></th>
<th>EBLA_H</th>
<th>EBLA_W</th>
<th>KULTEPE</th>
<th>NIPPUR</th>
<th>LARSA</th>
<th>UR</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBLA_H</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBLA_W</td>
<td>0.644</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KULTEPE</td>
<td>0.658</td>
<td>0.715</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIPPUR</td>
<td>0.561</td>
<td>0.684</td>
<td>0.685</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARSA</td>
<td>-0.116</td>
<td>0.236</td>
<td>-0.083</td>
<td>0.402</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>UR</td>
<td>0.617</td>
<td>0.708</td>
<td>0.854</td>
<td>0.88</td>
<td>0.211</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 5. Correlation Analysis of quantograms: scatter plots.
represented, although less relevant than the Mesopotamian one, and quanta in the range of 9.4 g (i.e., the “Levantine” unit) are detected at least in Kültepe and Ebla, but the substantial width of the peak of the Ur sample might also account for this unit. Finally, small peaks in the range of 6.8g-6.9 g might be related to a hypothetical “Aegean” unit or more probably to an independent system originally used for the wool in the Near East (Zaccagnini, 1999–2001). The simultaneous presence of balance weights belonging to different systems is a well-known fact, already pointed out in the respective publications of each sample included in our analysis (see § 4.1). The hacksilver sample follows exactly the same pattern: it can be observed that the hacksilver quantogram fairly matches several of the commonly accepted base units, and follows a very similar pattern to those related to the balance weights from Ebla and Kültepe. In other words, hacksilver shows clear signs of compliance with normative weight-systems.

5.2. Correlation analysis of quantograms

The hacksilver sample shows good correlations with almost every other dataset, with scores for r between 0.561 and 0.658 (Table 2, Fig. 5), and in particular with the balance weights from Ebla and Kültepe (0.644–0.658). It is interesting to point out that Ebla was the city in which the hoard was found and Kültepe was a commercial centre, certainly in direct contact with Ebla exactly during the period when the silver bullion was hoarded. High scores, beyond the significance level (i.e., 3.55; see Appendix 1), are also obtained between all couples, the only notable exception being the balance weights from Larsa. As already noted above, Larsa presents an unusual frequency distribution of weight values, that affects the results of the CQA.

In general, the results support a good overall correspondence between the Levantine and Mesopotamian weight-systems; furthermore, the hacksilver sample from Ebla always shows good correlations with other datasets, with the only exception of Larsa.

5.3. Frequency Distribution Analysis

The last question addresses the possibility to detect patterns that are not directly related to the compliance with normative systems. The analysis aims at understanding whether there are certain values that tend to occur significantly more often than others, and whether such frequent values can be compared across different distributions. The working hypothesis is that everyday practice and frequent commercial contacts can produce a convergence between different normative systems, towards a limited array of “Standard Average Quantities” (SAQ), that ultimately become more frequently used than others (Ialongo et al., 2018).

The cumulative frequency distribution of the complete dataset shows a series of extremely sharp “peaks”, each including significant concentrations from each of the analysed samples (Fig. 6. A). Interestingly, the average values of such peaks almost exactly correspond to multiples and submultiples of the Mesopotamian unit of 8.4 g, namely 1/2, 2/3, 1, 2, 3, 5, 10 and 20. The comparative distribution shows that the hacksilver sample produces clusters in the same value-ranges (Fig. 6. B). The vertical bands, illustrating the dispersion of each major peak at 1 standard deviation, include c. 2/3 of the total hacksilver sample: this means that, if we consider 2σ distributions, the peaks would account for nearly the total variability of the silver hoard.

Fig. 6. Frequency Distribution Analysis. A: cumulative distribution of the entire dataset; each peak is described by its mean value (on top) and by its Coefficient of Variation (below, in brackets). B: frequency distribution of the hacksilver sample from the Ebla hoard; the bands identify the dispersion of each peak in the cumulative distribution at 1 standard deviation.
6. Discussion

The analyses retrieve consistent results. CQA shows that the quantal configuration of the analysed samples is substantially in line with the expectations, showing that known units are represented everywhere. The correlation analysis appears to provide a solid base for comparison between quantograms, showing a high degree of compatibility between all the analysed samples, with the only exception of Larsa. In general, the results show that hacksilver has the same statistical properties of balance weights, and that fragmentation was probably aimed at obtaining predetermined weight values, i.e., the hacksilver from Ebla possesses the characteristics that one would expect from a bullion-currency. In particular, hacksilver shows clear signs of compliance with known official systems, and a high degree of affinity with weight-systems attested at Kültepe, Ebla, Ur and Nippur. Furthermore, FDA shows that balance weights and hacksilver from very distant contexts tend to cluster around the same quantities that are, in turn, strikingly close to logical multiples and submultiples of the alleged Mesopotamian unit of 8.4 g. To summarize: at local level, we observe that a multitude of normative units is responsible for the variability of both balance weights and hacksilver, but at a global scale everything appears much more related to the Mesopotamian unit than to any other normative system. We need to understand then, why different normative units tend to converge around the same approximate quantities.

An explanation based on historical considerations could be that the concentration of demand for commodities — not only for silver

![Graphic model of the convergence of different normative weight systems](image)
— in Mesopotamia was so prominent in the global market during the MBA, that a great deal of the overall transactions was carried out according to Mesopotamian weight-standards, or even directly by Mesopotamian agents in northern countries, as it is attested in Anatolia with the ports of trade settled by Assyrian merchants (Dercksen, 1999). This hypothesis can plausibly explain the global convergence towards the Mesopotamian standard, but it cannot explain the multitude of units that are observed locally. Are the “western systems” simply underrepresented, or can there be some other process at work, that “hides” them from our analyses? While maintaining the assumption that Southern Mesopotamia might have skewed the global market through the concentration of demand, we suggest that the “global convergence” can be explained without necessarily implying a large-scale use of the Mesopotamian “norm” in western and northern regions; more specifically, we expect that the statistical dispersion of the clusters observed in the FDA can largely account for the simultaneous presence of different normative systems.

Assuming that the alleged units for the MBA are correct — and excluding the supposed Aegean unit of 6.69, whose actual existence is questionable (Hafford, 2012) — we are dealing with basic values of 7.8 g, 8.4 g and 9.4 g, defining the “Syrian”, “Mesopotamian” and “Levantine” systems respectively. Such numbers are small and relatively similar to each other, but it can be argued that their relative difference was large enough to force traders to pay close attention to conversion factors. But how similar are they in practice, and how much dispersion exists between them? Exact units are pure numbers, and as such they possess a range of properties. The most interesting property for the problem at hand, is that big numbers can be divided for many different small ones, for a small error. This means that, while there can be a significant dispersion in converting different systems at the unit-level, the error becomes more and more negligible when one converts bigger multiples of the same units. For example, if we try to equate a shekel of 9.4 to one of 8.4 we obtain a remainder of 1, corresponding to an error of 11.9% with respect to 8.4 and of 10.6% for 9.4. It can be argued that this would be too much of a loss for a trader dealing in silver. But, if one is to trade 10 Mesopotamian shekels of silver (84 g), there is no need to employ a 1:1 ratio: one can just chose the closest multiple of the concurrent system, i.e., 9.4 g * 9 = 84.6 g for a remainder of just 0.6 g and an error of 0.7%, which is far smaller than the commonly accepted accuracy threshold of ±5%. Furthermore, it is vital to bear in mind that real-world measurements always produce some statistical dispersion, i.e., they are never factually exact. This means that when we are dealing with very close shekels the respective error distributions tend to overlap, to a point that it can often be impossible to discern between them (Hafford 2012): this is the case of the Syrian (7.8 g) and the Mesopotamian (8.4 g) units, whose respective error distributions largely overlap at 1 standard deviation of ±5%. Different systems of measurement are also formally “unified” around the standard value of the mina with a theoretical value of 470 g (the so-called ‘western mina’), widespread in the Levant alongside the Mesopotamian mina of 504 g (Parise, 1970–71, 1981, 1984; Ascalone and Peyronel, 2006: 23–25). The standard value of the ‘western mina’ was defined according to a ratio of 60, 50 or 40 units, characterizing, respectively, the Syrian, Levantine and Anatolian systems. Thus, 60 shekels ‘of Karkemish’, 50 shekels ‘of Ugarit’ or 40 shekels ‘of Khatti’ (with theoretical values of 7.8 g, 9.4 g and 11.75 g) were respectively required in order to obtain a mina of 470 g (Parise, 1984: 129).

Based on these simple observations, we can derive the following model: two or more different systems of measurement will tend to “get close” to each other as the magnitude of the quantity being measured increases, to a point where the error distributions of the respective multiples will largely overlap. In order to illustrate the model, a purely theoretical “meta-system” was calculated, based on the average values of the closest multiples of the three shekels (Fig. 7. A). The figure shows the relative difference of the three systems from the meta-system, the latter corresponding to the zero-line: the overall relative error between the exact values of the multiples of the different shekels tends to decrease exponentially, steadily falling below 5% starting around 40–50 g, i.e., around factors of 4–5 of each shekel. Such a model provides the framework to test our expectation (Fig. 7. B): if we plot the means and standard deviations of the clusters produced by FDA and we compare them to the closest multiples of the Syrian and Levantine shekels (also plotted with a standard deviation of ±5%) we observe that the Syrian system always overlaps the clusters’ dispersion, that the Levantine system does so at the values of 2/3 and of 4 shekels (≈ 5 Mesopotamian shekels), and that all three systems become substantially indistinguishable starting at around 10 Mesopotamian shekels (≈ 11 Syrian shekels = 9 Levantine shekels).

Both hacksilver and balance-weights from different normative systems appear to converge towards “Standard Average Quantities” that produce the least possible statistical dispersion when converted into different unit-systems. Given its nature as a standard medium of exchange, and its ubiquitous circulation, hacksilver is most probably to be acknowledged as the medium of the practical convergence of different normative systems. Such a convergence effectively produces a “meta-system”, aiding cross-system conversion, and thus international exchange. At the same time, the meta-system is also very similar to the Mesopotamian one: this suggests that the Southern Kingdoms, in the MBA, might have

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1.05</td>
</tr>
<tr>
<td>2.10</td>
</tr>
<tr>
<td>3.15</td>
</tr>
<tr>
<td>4.20</td>
</tr>
<tr>
<td>5.25</td>
</tr>
<tr>
<td>6.30</td>
</tr>
<tr>
<td>9.45</td>
</tr>
<tr>
<td>10.50</td>
</tr>
<tr>
<td>12.60</td>
</tr>
<tr>
<td>15.75</td>
</tr>
<tr>
<td>21.00</td>
</tr>
<tr>
<td>31.50</td>
</tr>
<tr>
<td>52.50</td>
</tr>
</tbody>
</table>

Fig. 8. Scores of the coefficient r (y axis) for the correlation between the ideal dataset and a test dataset. The analysis is reiterated 100 times (x axis), with the test dataset increased by 0.0015 for each iteration.
Fig. 9. Correlation Analysis of quantograms: randomly generated test-datasets. A: values of $r$ for the correlation between the ideal dataset (Table 1) and two test datasets; test datasets are randomly generated based on the ideal dataset, with each value having equal probability within a range of 0.5 and 2.0 from the corresponding value in the ideal dataset; the analysis on each couple is reiterated 100 times. B: values of $r$ for the correlation between two test datasets; each value in each dataset has an equal probability within a range of ±5% from the corresponding value in the ideal dataset. C: values of $r$ for the correlation between five couples of test datasets; each value in each dataset has an equal probability within a range of either ±5% or ±10% from the corresponding value in the ideal dataset.
expressed a demand for commodities — for silver as well as for other goods — that was so prominent in the global market to shape an international system based on round multiples of their own normative units.

7. Conclusions and perspectives

Based on the results of different statistical analyses, we conclude that the starting hypothesis is valid, and that hacksilver currency in the Ancient Near East was shaped and/or fragmented in order to comply with the normative weight-systems in use in the trade network where it circulated.

Further patterns are highlighted by the analyses, which can be delved upon in order to formulate new hypothesis, to be tested in future research. Our results suggest that the wide diffusion of a standard bullion-currency can effectively bypass the constraints imposed by the different normative-systems in use in a given trade network, while still complying with each of them. In other words, the existence of different official weight-systems does not constrain international trade in any substantial way. Economic practice can produce meta-systems, providing acceptable accuracy in the conversion between different standards, while still complying with each official system (Fig. 7). The existence of an “internationally” acknowledged bullion-currency must have played a key-role in the emergence of meta-systems in a network, providing an objective, shared frame of reference. As a consequence, the degree of affinity between different systems should be proportional to the intensity of interconnections within trade networks, and observable through the quantitative analysis of balance weights and weight-regulated bullion-currencies.

Acknowledgements

Funding: This work was supported by the European Research Council under the European Union’s Horizon 2020 Framework Programme and was carried out within the scope of the ERC-2014-CoG ‘WEIGHTANDVALUE: Weight metrology and its economic and social impact on Bronze Age Europe, West and South Asia’, based at the Georg-August-Universität of Göttingen, Germany [Grant no. 648055]; Sapienza University of Rome, Department of Sciences of Antiquity; IULM University of Milan, Department of Humanities. We wish to express our gratitude to Sarah Clegg for the constructive critiques, and to the anonymous reviewers for their precious advice.

A. Appendix 1

Several sets of ideal and randomly generated datasets were analysed through CQA, and the similarity of the resulting quangograms was tested through Pearson’s product-moment correlation (r). As a base value-list for all our tests we use an arbitrarily defined dataset of 13 ideally exact multiples of 1.05; the value-list has no particular meaning, other than being the same employed by Petruzzo (1992: 74) in his own tests (Table 3).

The first test involves comparing two exact series, the first one remaining the same and the second one progressively increased by 0.0015 in each subsequent iteration. The aim is to assess the response of CQA in analyzing datasets that are similarly “quantally configured” (i.e., whose values are multiples of a similar array of basic units), and to set an ideal threshold beyond which two datasets can be considered “different”. Pearson’s r was computed for 100 iterations, for 100 quanta with an increment of 0.0015 (Fig. 8). The curve is very steep in the first part, and stabilizes around zero starting from a relative difference of c.a. 8%, thus indicating that small variations strongly affect the similarity between quangograms.

The second test aims at exploring the correlation coefficients for two random, non-quantal datasets, in order to define the expectations for chance results. Two samples were randomly generated, starting from the ideal value-list; in order to obtain samples with different values but similar distributions, each entry in the list was added a random fraction of ±50%. The ideal value-list and the two randomly generated samples were then compared, and Pearson’s r was computed for 100 iterations for each couple, and for 100 quanta with an increment of 0.1, starting from 1.0. 95% of results give values of r between −0.388 and 0.355 (mean = −0.029, 2r = 0.407) (Fig. 9, A). The upper value of the 95% interval (0.355) is assumed as a significance level, beyond which values for r are very unlikely to result by chance.

The third test is intended to approximate a real research-case, by simulating samples with controlled dispersion. The aim is to observe whether the results for r are significantly higher than the 95% significance level computed for random samples (r = 0.355).

Four datasets were generated, starting from the ideal value-list: two datasets were obtained by adding a random fraction of ±5% to the original values, and two more by adding a random fraction of ±10%. Pearson’s r was computed for 100 iterations for each couple, for 100 quanta with an increment of 0.1. The correlation between the two samples with ±5% dispersion is consistently strong, with a 95% interval for r between 0.494 and 0.935 (mean = 0.715, 2r = 0.221) (Fig. 9, B), and always higher than the threshold of 0.355. The remaining couples show consistently lower values of r, with a 95% interval between 0.234 and 0.768 (mean = 0.501, 2r = 0.267) (Fig. 9, C); 86.2% of the results are above the 95% significance level of 0.355.

The results of the preliminary tests show that Pearson’s r is a good overall measure for the similarity between quangograms. For values lower than 0.355 r is not significant, and all results below this threshold can be due to chance. For values higher than 0.355 r is significant, and the strength of the correlation is proportional to the similarity of the quanta employed in the calculation of test-datasets. High values for r, therefore, probably indicate that the quantal bases of the analysed datasets have a difference of ±5% or less, as suggested by the first and third tests (Figs. 8 and 9, B).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jas.2018.01.002.

References


Alberti, M.E., Parise, N.F., 2005. Towards a uni...
