Where do measurement units come from?

Kensy Cooperrider (kensy.cooperrider@northwestern.edu)
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

Dedre Gentner (gentner@northwestern.edu)
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

Abstract

Units as they exist today are highly abstract. Meters, miles, and other modern measures have no obvious basis in concrete phenomena and can apply to anything, anywhere. We show here, however, that units have not always been this way. Focusing on length, we first analyze the origins of length units in the Oxford English Dictionary; next, we review ethnographic observations about length measurement in 111 cultures. Our survey shows that length units have overwhelmingly come from concrete sources—body parts, artifacts, and other tangible phenomena—and are often tied to particular contexts. We next propose a reconstruction of how abstract units might have emerged gradually over cultural time through processes of comparison. Evidence from how children understand length and measurement provides support for this account. The case of units offers a powerful illustration of how some of our most important, pervasive abstractions can arise from decidedly concrete, often embodied origins.

Keywords: measurement; abstraction; units; comparison; language; culture; analogy

Introduction

Poppy seeds and barley-corns, poles and bows, chain links and goads for driving oxen. For English speakers in earlier times, these were tools of measurement. Along with spans of the body—including finger-widths, hand-breadths, and arm-lengths—such objects furnished methods of measurement that were highly concrete and often context-specific. For centuries in Britain and across Europe, certain units were favored for measuring cloth, others for horses, others for land; different units were used for length, depth, and distance; and these units often varied from one town to the next (Kula, 1986). This stands in stark contrast to measurement today. Our main units are now highly abstract. That is, many have no obvious tie to concrete objects or to the body, and they may be used for measuring anything, anywhere. Indeed, in the International System of Units, the measurement of length, area, and volume are all based on a single universal unit—the meter—and its derivatives (e.g., millimeters, kilometers, etc.). And this lone unit is now officially defined, not with reference to the body or any tangible object, but in terms of the distance light travels in a fraction of a second.

Our purpose here is to account for this shift in the nature of measurement, from highly concrete to highly abstract—to ask how we got from measuring cloth against our elbows and distance with our feet to measuring everything against a universal abstraction. Some parts of this shift have happened only recently. Over the last two centuries, countless experiments have made possible the precision with which our current units are now defined, and international agreements have spread the metric system across the globe (Alder, 2002; Crease, 2011). But long before scientists took up the task of making units more precise, more elegant, and more systematic, and long before the metric system was even conceived, humans had to develop units in the first place. Our survey suggests that units evolved gradually; indeed, units in the modern sense may not even have appeared in all cultures. How did they emerge? Here, we draw on data from a variety of sources—linguistic, ethnographic, historical, and developmental—to try to reconstruct the evolution of units.

The foundation of our proposal is that abstract units emerge slowly out of processes of *comparison*—comparisons that are initially between physical objects. Indeed, some have claimed that the "real essence of measurement is comparison" (Crump 1990, p. 77). The importance of comparison in fostering learning is now recognized. For instance, prompting learners to compare examples helps them arrive at new abstractions (e.g., Alfieri, Nokes-Malach, & Schunn, 2013; Gick & Holyoak, 1983; Kurtz, Miao, & Gentner, 2001). The importance of comparison in the formation of knowledge is also evident on historical timescales. Many abstract concepts started out as novel metaphors—that is, one-off figurative comparisons—and then became conventional (Bowdle & Gentner, 2005). Such abstractions sometimes enter the lexicon (Xu, Malt, & Srinivasan, 2017), allowing people to use them without necessarily recognizing where they came from. A similar general trajectory, we hypothesize, is evident in the case of measurement. To preview, we reconstruct this trajectory as involving overlapping stages. First, people make ad hoc comparisons between a concrete thing—a *target* to be measured—and another concrete thing—a *comparator* (e.g., an actual foot). Second, people come to favor certain comparators over others—in short, conventions emerge. Third, people abstract across these conventional comparators (e.g., many examples of feet) to develop an idealized, standardized comparator (e.g., the foot). Only at this stage does measurement involve anything resembling our modern notion of a fully abstract unit. Later

stages also involve comparison, along with cognitive processes such as proportional reasoning.

In reconstructing this emergence, we focus on length—including other measures of linear extent such as height, depth, and distance. However, our proposal should apply in broad strokes to other basic physical dimensions, such as weight and volume. We consider three interwoven aspects of length measurement: what linguistic resources people have for talking about length (e.g., unit terms, constructions); what tools and practices people use to measure length; and what cognitive representations people have of length and of length units. We start by examining the evidence that measurement concepts began as highly concrete, not only in Europe but across the globe. This examination lays the groundwork for important generalizations about how measurement concepts could have emerged. We then present our proposed reconstruction of the critical stages in the shift from highly concrete beginnings to our current system. Finally, we note parallels between how units emerged over historical time and how they emerge in child development. We begin with an examination of the sources of unit terms in English.

Length units in English

The examples already presented give some flavor of the concrete basis of measurement units in English. But just how common are concrete sources for measurement units, relative to abstract ones? To answer this question, we examined all the linear measures in the Oxford English Dictionary (OED) (http://www.oed.com/). The OED includes 53 words for length measures (excluding recent borrowings), 48 with known origins. We classified the sources of these words into concrete and abstract, and further subdivided the concrete sources into body-based, artifact-based, and other. We discuss each category in turn.

Concrete (body-based). The most frequent concrete source for length units in English is the human body, accounting for 33% of the words (16 of 48). Beyond the familiar 'foot,' these terms include the 'fathom' (the length between the outstretched arms), the 'ell' (the full length of one arm), and the 'cubit' (the span from the elbow to the fingertips). Shorter spans include the 'hand' and 'palm,' and terms based on both finger length and finger breadth.

Concrete (artifact-based). Length terms derived from artifacts account for another 29% (14 of 48). These include 'yard' (originally a type of pole), and other terms for similarly elongated objects, such as 'rod,' 'perch,' and 'virgate.' Other artifact-based terms include 'bow,' 'chain,' 'link' (of a chain), and 'goad,' a tool used for driving oxen.

Concrete (other). Terms deriving from other concrete sources account for a further 23% (11 of 48). These include 'poppy seed,' 'barley-corn,' 'oxgang' (based on amount of land plowable by a team of oxen), 'reed,' and 'furlong' (a compression of 'furrow' + 'long', based on the length of a furrow in a field).

Abstract. The remaining 15% (7 of 48) have an abstract origin. These include 'inch,' from the Latin for 'twelfth,' 'mile,' from the Latin for 'thousand,' and 'meter,' which derives from a Greek word for 'measure.'

In sum, while measurement terms in English are not inevitably concrete in origin, they are overwhelmingly so. The OED analysis also provides evidence for another sense in which measurement terms are concrete: many of these units were primarily used in specific contexts. 'Bow' was confined to archery; 'chain' and 'prime' were used chiefly in surveying; 'furrow' and 'land' were used in agriculture.

The etymology of English unit terms provides a valuable window into the history of measurement, but it only gets us so far. For one, it only provides that window for one culture. Moreover, a dictionary does not always offer documentation of the practices associated with these units. To address these limitations, we next cast a much wider net, looking across cultures and across historical eras at measurement terms and practices that have been documented by ethnographers and historians.

Length measurement across cultures

To understand the course of measurement units, we need to go beyond world languages like English (Lupyan & Dale, 2010; Majid & Levinson, 2010) and also examine practices in other cultures. To do this, we analyzed the Human Relations Area Files (HRAF) database, which compiles and topic-codes ethnographic accounts from 311 cultures, across every geographic region (http://hraf.yale.edu/). The database includes observations about the topic of 'Weights and measures' from 193 cultures; from this set we identified 114 cultures for which observations about length measurement were available. We supplemented this database with in-depth studies of measurement within particular cultures (e.g., Alkire, 1970; Hallowell, 1942; Pankhurst, 1969), and with general overviews of the history of measurement (e.g., Alder, 2002; Crease, 2011; Kula, 1986).

A first important conclusion from this survey is that not every culture has developed units as we understand the term. Moreover, even where units are reported, researchers often comment on their "semi-standardized" (Hallowell, 1942) nature, or qualify them as not being "mathematical" (Richards, 1939) or "precise" (Best, 1919). With these caveats in mind, we characterize trends in the units reported for small-, medium-, and large-scale extents.

Small-scale extent. The units described for measuring small extents are overwhelmingly based on the body (Table 1). Best (1919) describes 13 body-derived units used by the Maori, from one equivalent to the first joint of the thumb to one meaning ten fathoms. Alkire (1970) lists 14 body-derived units used in the Caroline Islands, from as short as a finger joint to as long as a fathom. The Tzeltal of Mexico had a series of units, from the *nab*, the span between the thumb-tip and end of middle finger when all fingers are extended, to the *yankabal*, the distance between the armpit and the fingertips of the opposite arm when outstretched (Villa Rojas, 1969). Such examples of body-based units are found worldwide. Across these systems, certain anatomical spans predominate, particularly ones based on

Table 1: Concrete sources of measurement units across cultures

Source	Source type	Scale	Culture/ Region
Louse egg	Other (natural world)	Small	India
Barley-corn	Other (natural world)	Small	Anglo-Saxon
Finger (breadth)	Body-based	Small	*widespread*
Hand span (thumb to little finger)	Body-based	Small	*widespread*
Foot	Body-based	Small	*widespread*
Pace	Other (activity)	Small	*widespread*
Forearm (elbow to fingertip)	Body-based	Small	*widespread*
Fathom (wing-span)	Body-based	Small	*widespread*
Person (with arms extended above head)	Body-based	Small	Maori
Furrow (length of furrow in a field)	Other (agricultural)	Medium	Anglo-Saxon
Hatchet throw, backwards from seated position	Other (activity)	Medium	Hungary
Hatchet throw, from seated position	Other (activity)	Medium	Hungary
Bow shot	Other (activity)	Medium	Kogi
Call (distance from which can be heard)	Other (activity)	Medium	India
Distance seen from camel's back	Other (activity)	Large	Saharan Africa
Day (distance covered in day of walking)	Other (event)	Large	*widespread*

salient divisions of the forelimbs. Less often reported are examples of measures based on anything *other* than the body. However, several are based on artifacts, such as a unit of length based on a long tool for cutting banana leaves, used by the Chagga (Africa) (Marealle, 1963).

Medium-scale extent. Traditional measures for medium-scale extents are less widely reported. They tend to involve brief actions. Examples include 'bow shot' and 'stone throw.' In other places, sound served as the basis, with measures derived from the distance at which one could hear a person calling or a bull bellowing. Elsewhere, sight served as a basis, with measures based on the distance one could see from a camel, in the Sahara (Kula, 1986), or the distance at which a buffalo appears to be the size of a man, in Vietnam (Pasquier, 1907).

Large-scale extent. At larger scales, measures are predominantly based on temporally extended events. In culture after culture, large distances are reckoned in terms of days spent walking, or 'sleeps.' Some measured the distance of journeys in terms of culture-specific consumption habits, such as the number of betel nuts chewed (Karen), pipe-bowls smoked (Ojibwa), or young coconuts drunk (Nicobarese). At least one culture, the Ojibwa of North America, found an ingenious way to use the body to measure large-scale distances. This was done by superimposing the outstretched hand on the arc of the sun. One 'hand-stretch' was considered one fourth of the arc from sunrise to zenith, and could thus be used to estimate how much of a day it would take to travel the target distance (Jenness, 1935).

Many measurement practices in traditional societies are described as confined to particular contexts (Table 2). The Kedang of Indonesia, for instance, have a conventional set of points on the arm used for measuring wedding gifts—chiefly, elephant tusks and gongs (Barnes, 1982). Among the Siwai of Papua New Guinea, a conventional system arose for measuring the girth of pigs (Oliver, 1955). A practice for measuring string money among the Yurok in California sometimes involved tattooing measurement landmarks on the arm (Kroeber, 1925).

Table 2: Context-specific measurement practices across cultures

Target	Comparator	Culture
String money	Forearm tattoos	Yurok
Buffalo horns	Forearm points	Toraja
Canoes	Hand points	Caroline Islands
Wedding gifts (gongs, tusks)	Forearm points	Kedang
Pigs (girth)	Arm points	Siwai

Reconstructing the emergence of units

How could abstract units as we know them today have emerged out of such highly concrete, often context-specific practices? Our proposal takes off from the observation that comparison is the basis of measurement (e.g., Crump, 1990). In a sense, the practical activity of aligning an object against a ruler or other tool can be seen as a comparison.

But, we argue, the role of comparison in measurement also goes much deeper. Though the power of comparison in learning is increasingly recognized, its power in cultural history is less appreciated. Comparison is claimed to be universal across cultures (Brown, 1991), and enters into language and cognition in several ways. Only a subset of comparisons concern physical properties like length, and even these are very often relative and qualitative (e.g., "This log is longer than that one"). But such comparisons are not measurement in a strict sense (Hallowell, 1942). The impulse to measure is an impulse to express, more or less precisely, *exactly* how long a target is. We suggest that, initially, this impulse motivates ad hoc comparisons between particular objects—a target and another object—and, later, comparisons between other things.

Stage 1: Ad hoc comparison

The first type of comparisons that enter into the early stages of measurement are *ad hoc*. A person has some target object—a log, fish, or house—whose length they would like to characterize. To do so, they invoke a point of comparison—a *comparator*—whose length is more accessible contextually or is more widely known. There are a number of contexts where such ad hoc comparisons would have been used—and, indeed, are still used. One is when someone is trying to characterize the properties of a non-present target, such as a fish that got away. Invoking this target can be done purely in language, using a more widely known object as a comparator, as in "It was the length of my arm." It can also be done by anchoring the comparison to the present context, e.g., "It was this big," accompanied by a size gesture. Gesture regularly enters into such comparisons, and gestural conventions specialized for ad hoc comparison have been widely reported (e.g., in Nuer; Huffman, 1931).

Another prominent context for ad hoc comparison is when one is trying to judge which of two targets, A or B, is longer. This can be done by eye when the difference is marked, or by directly juxtaposing the targets when this is possible. However, when A and B cannot be directly juxtaposed, the judgment requires a new solution (Hallowell, 1942). Consider the example of wanting to know which of two spatially separated trees has a thicker trunk. If "eyeballing" is unsatisfactory, a solution is to introduce a comparator—a third object that can be directly juxtaposed with each target. This comparator could be a body part, tool, or something improvised on the spot to match one of the two targets—for example, a twig broken off or a cord cut to the right length. Techniques like this are widely described in the ethnographic literature (e.g., Hallowell, 1942), and can be readily observed today. In the game of bocce, for instance, when one wants to decide which of two balls is closer to the pallino, a string is first used to mark the distance of one ball (creating a comparator to match one target length), and then used to compare this to the second target length. Interestingly, the bulk of these ad hoc comparisons seem to occur between a target and a comparator of equal length (*equal comparisons*), such that no further computation is involved, rather than between a target that is shorter or larger than the comparator (*unequal comparisons*).

Stage 2: Conventional comparisons

Over time, certain comparators that were initially used on an ad hoc basis become conventional. That is, rather than reach for a comparator simply because it is apt for present purposes, people begin to use comparators that are commonly used as such in their community. Such conventionalization processes are evident in language. A recent analysis tracked changes over time in how English speakers have filled the construction "the size of a [NOUN]," using the Google Ngram corpus (Morris, 2017). In the 1800's, the top four nouns in this construction were *pea*, *walnut*, *pinhead*, and *egg*; in the 2000's, the top four have been *pea*, *walnut*, *quarter*, and *football field*. This analysis shows that certain comparators (*pea*, *walnut)* have stability over time, whereas others (*quarter, football field*) are newly conventionalized.

Why do some length comparators become conventional, rather than others? This could reflect factors such as *availability* (i.e., how readily available the comparator is—a barley-corn would never catch on outside of a farming community), *alignability* (how readily the comparator can be aligned with the target—the human ear is readily available, but is not easy to align with a target), and *aptness of scale* (the foot is available and alignable, but would be impractical when measuring distance between towns) (Crease, 2011). The universal use of body-based spans for measuring small-scale extents is perhaps explainable in terms of these factors: the body is always available, readily alignable (certain spans more than others), and apt for small-scale extents.

Conventionalization appears to be a gradual process, and we suggest that a few changes happen as conventionalization proceeds. For one, people become more likely to use comparators in *unequal* comparisons. Unequal comparisons are more cognitively taxing, requiring either proportional reasoning (e.g., half the length of foot) or counting (the length of three feet). Researchers have noted that, in a given community, only certain comparators enter into unequal comparisons. In describing body-based units in the Caroline Islands, Alkire (1970) notes that, while one can speak of "two forearm-lengths" or "two hand-spans," informants reject the same construction with other spans, such as palm-width. Instead, one must say something like "two spans of palm-width size" as such terms are only used to describe "the exact length of an object relative to any other" (p. 29). We interpret this as evidence that these spans differ in their degree of conventionalization as length comparators.

Stage 3: Standardized comparisons

Once comparators have become conventional, people begin to recognize a problem: though a comparator such as the

foot is conventional, not all instances of that comparator are strictly equivalent. This leads to imprecision, and the historical and ethnographic literature is full of evidence of people recognizing—and trying to work around—this issue. At one point in China, a distinction was made between a 'man's hand' and 'woman's hand' (Crease, 2011). Among the Maori, it was common to enshrine one man's arm span in a wooden rod, called a *rauru*; such rods would be used throughout house-building projects and would sometimes be passed down through generations (Best, 1918, p. 31). Elsewhere, people have taken advantage of the fundamental imprecision of body-based measures. In Ethiopia, it was common to bring to the market a person with "long arms" to help one measure purchases (Pankhurst, 1969, p. 36).

A solution to the problem is to develop an idealized version of the conventional comparator, or *standard*. Though the issue of imprecision is likely widely recognized in traditional societies (e.g., Saxe & Moylan, 1982), few appear to have introduced standards. These emerge with intensive commerce. Standards could be based on known concrete lengths (e.g., a prestigious person's arm), or they could be devised to be representative of the class of comparators. Often the standards nod to their concrete origins, as when King Edward I (1239-1307) introduced the "Iron Ulna" as a standard measure. Importantly, we suggest, the adoption of a standard precipitates a key change in how the comparator will come to be conceived. Decoupled from any actual object—or even a class of objects—the comparator becomes something significantly more abstract and more recognizable as a unit.

Standardization may precipitate other changes, too. For instance, we speculate that it is at this stage that the comparator becomes especially likely to break free of its use in particular contexts. Further, it is only at this stage that it makes sense for people to propose new definitions of the standard. History is full of such proposals. Once the foot was not only an appendage but an idea, it made sense for John Locke to propose the "philosophical foot" and for others to introduce the "natural foot" and even the "manual foot" (Whitehall, 2007). Moreover, once the idea of abstract standards is in play, there is no need to simply redefine (formerly) concrete measures; one can invent entirely new standards, abstract form birth. The meter, first proposed in 1670, is the pre-eminent example.

Later stages

After the notion of a standard—and with it the idea of an abstract unit—is established, the stage is set for further changes, which we sketch only very briefly. Comparison is involved in these further changes, too, but alongside other cognitive processes. The next key change is toward *systematization*. This occurs when people begin to compare standard units to each other, abstractions to abstractions. This may first happen within the same dimension, when people realize the utility of having units neatly nested within other units. This process likely overlaps with standardization. Since an abstract unit allows wiggle room for redefinition (as in the different versions of the "foot" already mentioned), it can be redefined to allow for hierarchical relations with other units (Kula, 1986). Systematization of this type eventually also happens across scales and dimensions. That is, instead of merely comparing two different units of small-scale extent (e.g., 'foot' and 'yard'), people start to compare length units at different scales (e.g., 'foot' and 'mile'); and then units of length with units of depth, and so on. The process depends, not just on comparison, of course, but also on proportional and hierarchical reasoning. The end result of this process of systematization is a compact, coherent system of units that can be understood entirely *relationally*. Millions of people today use the meter without knowing—or needing to know—its physical basis in the speed of light.

Length measurement in child development

We have proposed that length units emerge gradually, in stages. Do children's ideas about length follow a similar course? Several sources of evidence suggest this may be the case. A first source is studies by Piaget and colleagues on "spontaneous measurement" in young children (Piaget, Inhelder, & Szeminska, 1960). These researchers examined what children would do when asked to judge which of two towers is taller but could not physically juxtapose the two. Several stages in children's behavior were evident. Until the age of four or five, children were content with rough, visual comparisons. Later they spontaneously attempted to use a comparator—whether their own body or a stick. At first children only considered such a stick useful if it was the same length as one of the targets (i.e., one that afforded an equal comparison); only later, not until the age of 7 or 8, did they use a comparator longer or shorter than either target (i.e., ones that afforded only unequal comparisons). Follow-up studies on spontaneous measurement find that children can be induced to use a comparator at younger ages if the inadequacy of visual comparison is more obvious (Bryant & Kopytynska, 1976), or when the task is couched in a particular, motivating context (Miller, 1989). Interestingly, children induced to measure in one context will not necessarily spontaneously measure later, in a superficially different context (Bryant & Kopytynska, 1976). These findings resonate with the cross-cultural and historical evidence reviewed above that length units often emerge within—and remain confined to—specific contexts.

Another line of research examines how children notice and remember length. This work shows that, in infancy, children are sensitive to *relative* length with a comparator present (i.e., how long a target dowel is relative to an adjacent dowel) (e.g., Duffy, Huttenlocher, Levine, & Duffy, 2005), but are not sensitive to *absolute* length until many years later. These researchers suggest that the ability to encode absolute length may result from having internalized a standard comparator that one can impose mentally on the target (Duffy, Huttenlocher, & Levine, 2005; Vasilyeva, Duffy, & Huttenlocher, 2007). In sum, measurement understanding is slow to emerge, both

developmentally and culturally. On both time scales, it may proceed through predictable stages and have to overcome context-specificity; and, speculatively, on both time scales, acquiring new measurement practices may scaffold new cognitive representations.

Conclusion

Units as we now know them—abstract, elegant, general, systematic, relationally understood—have not always been that way. They started out concrete, imprecise, bound to particular places and practices. Indeed, the very idea of abstract units, we suggest, is a cognitive and cultural achievement. It is easy to overlook this. Educated people in industrialized societies are so accustomed to parsing the world in terms of units—not just for length, but also weight, time, temperature, and many other dimensions—that we see the idea of units as self-evident. Our account suggests that this is likely mistaken: measurement units do not come easily. Despite having decidedly down-to-earth origins, measurement units—perhaps like numbers, spatial prepositions, cardinal directions, and other concepts—have now become so abstract and so ubiquitous that it is easy to forget they have a history at all.

References

Alder, K. (2002). *The measure of all things: The seven-year odyssey and hidden error that transformed the world*. New York: Simon & Schuster.

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87–113.

Alkire, W. H. (1970). Systems of measurement on Woleai atoll, Caroline Islands. *Anthropos*, *65*(1970), 1–73.

Barnes, R. H. (1982). Number and number use in Kedang, Indonesia. *Man*, *17*(1), 1–22.

Best, E. (1918). The Maori system of measurement. *The New Zealand Journal of Science and Technology*, *1*, 26–32.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216.

Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.

Bryant, P. E., & Kopytynska, H. (1976). Spontaneous measurement by young children. *Nature*, *260*(29 April), 773.

Crease, R. (2011). *World in the balance: The historic quest for an absolute system of measurement*. New York: W. W. Norton.

Crump, T. (1990). *The anthropology of numbers*. New York: Cambridge University Press.

Duffy, S., Huttenlocher, J., Levine, S., & Duffy, R. (2005). How infants encode spatial extent. *Infancy*, *8*(1), 81–90.

Duffy, S., Huttenlocher, J., & Levine, S. (2005). It is all relative: How young children encode extent. *Journal of Cognition and Development*, *6*(1), 51–63.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38.

Hallowell, A. I. (1942). Some psychological aspects of measurement among the Saulteaux. *American Anthropologist*, *44*(1), 62–77.

Huffman, R. (1931). *Nuer customs and folklore*. London: International Institute of African Language and Culture. Retrieved from http://ehrafworldcultures.yale.edu/document?id=fj22-014

Kroeber, A. L (1925). *Handbook of the Indians of California*. *Bulletin*. Washington: Government Printing Office. Retrieved from http://ehrafworldcultures.yale.edu/document?id=ns31-009

Kula, W. (1986). *Measures and men*. Princeton, NJ: Princeton University Press.

Kurtz, K. J., Miao, C.-H., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences*, *10*(4), 417–446.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS One*, *5*(1), e8559. http://doi.org/10.1371/journal.pone.0008559

Majid, A., & Levinson, S. C. (2010). WEIRD languages have misled us, too. *Behavioral and Brain Sciences*, *33*(2-3), 103-103.

Marealle, D. (1963). Notes on Chagga customs. *Tanganyika Notes And Records*, (60), 67–90.

Miller, K. F. (1989). Measurement as a tool for thought: The role of measuring procedures in children's understanding of quantitative invariance. *Developmental Psychology*, *25*(4), 589–600.

Morris, C. (2017, October 30). The size of things: An ngram experiment [Blog post]. Retrieved from http://colinmorris.github.io/blog/size-of-things

Oliver, D. L. (1955). *Solomon Island society: Kinship and leadership among the Siuai of Bougainville*. Cambridge: Harvard University Press.

Pankhurst, R. (1969). A preliminary history of Ethiopian measures, weights, and values. *Journal of Ethiopian Studies*, *7*(1), 31–54.

Pasquier, P. (1907). *Annam of former times*. Paris: Augustin Challamel. Retrieved from http://ehrafworldcultures.yale.edu/document?id=am11-048

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry*. New York: W. W. Norton.

Saxe, G. B., & Moylan, T. (1982). The development of measurement operations among the Oksapmin of Papua New Guinea. *Child Development*, *53*(5), 1242.

Vasilyeva, M., Duffy, S., & Huttenlocher, J. (2007). Developmental changes in the use of absolute and relative information: the case of spatial extent. *Journal of Cognition and Development*, *8*(4), 455–471.

Whitelaw, I. (2007). *A measure of all things: The story of man and measurement*. New York: St. Martin's Press.

Xu, Y., Malt, B. C., & Srinivasan, M. (2017). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. *Cognitive Psychology*, *96*, 41–53.