Heuristics in Judgment Tasks with Unrecognized Elements:
Supplementary Material

Heuristics Without Recognition

Anthroponymy

In order to examine the plausibility that the surnames included cues (such as the ending –ov) that the subjects may have used to aid them with their inferences, I resorted to anthroponymy. In anthroponymic studies, the term personal name (or simply name) is used sometimes to refer to forenames without a surname, sometimes to surnames without a forename, and yet other times to both elements. I will henceforth use the term anthroponym to refer indistinctly to any of the forms adopted by a personal name; name will refer to forenames only, and surname will denote parts of personal names such as the ones included in the lists of chess players.

Faure Sabater, Ribes Lafoz, and García Sancho (2002) notes that surnames were born out of the need to establish nominal distinctions between individuals when forenames began —for different reasons— to repeat, thus facilitating heritage distribution, maintenance of legal ownership records, etc. Their linguistic structure was varied in origin: a surname could allude to the person’s genre, some physical trait, his provenance, occupation, social rung, etc. It wasn’t until several centuries later that official regulations dictated the mandatory nature of surnames, barring the creation of new ones, establishing the rules for heir designation, etc. These regulations, combined with cultural traditions, gradually favored the creation and replication of certain regularities (patterns or repetitions) found in many names and surnames of today. One of the most widespread traditions consisted in giving newborns the same name or surname of one of their parents with the addition of different morphemes. When a morpheme is added in front of the parent’s name it is called a prefix; when appended at the end, a suffix; for any other position
it is given the name of *infix*. A morpheme is the smallest analyzable unit that is grammatically meaningful. There are bound morphemes, which appear as part of a word (e.g., in English, prefixes such as *un-* and suffixes such as *–ly*; in Spanish, prefixes such as *des-* and suffixes such as *-mente*), and free morphemes, which appear in isolation (e.g., *and, the, of* in English; *y, el, de* in Spanish). Bound morphemes change the meaning of a word by modifying its lexeme or root. As an example, *friend* is the root of the words *friends* and *friendly*; the morpheme *-s* denotes the plural form, while *-ly* extends the meaning to signal a trait or way of doing things. Both of these morphemes (which are suffixes, due to their position in a word) are regularly used in English to modify words so that they convey number and manner. Faure Sabater et al. point out that, in the case of anthroponyms, certain morphemes have also become regularities by following relatively stable conventions pertaining to their culture, language or region of origin, as well as to certain characteristics of the individual being named (their condition as sons/daughters, their gender). For example, the suffix *-ez* in Spanish was the most common among the suffixes that meant “son of”; thus, Martínez became the surname of the son of someone named Martín (*Martín + ez*). A man named Ramiro would call his son Ramírez. In Irish surnames such as O'Donnell, the prefix *O* became the contracted form of *of*, used to designate the descendants of a man named Donnell. The prefixes *ibn, bar* and *ben* (reproduced here in Latin characters) were commonly used in Semitic languages such as Arabic or Hebrew to mean “son of”. As mentioned in the article, Álvarez (1968) points out that Slavic surnames also have characteristic suffixes that make them identifiable. In Russia, for instance, adding the suffix *-ov* to a son’s name was habitual practice. In Poland, *-sky* or *–skiy* are common, while *-ko, -nko*, and *-enko* are typical of Ukrainie. These endings vary according to the Slavic region, country or language where the names originated. However, a country can have many characteristic suffixes (e.g., in Russia the suffixes *–in* and *–ich* are also common), and suffixes tend to spread to other countries over time (*-ov* is also common in surnames from Bulgaria, Belarus and Serbia). The lack of a one-to-one correspondence between suffixes and nationalities seems to
pose an unsurmountable barrier for inferring one from the other, but later I will detail the specific
cognitive abilities that allow solving these and similar problems.

Detection, Recording and Retrieval of Anthroponyms

With the object of examining the plausibility of the subjects deeming certain surnames as Russian or Slavic, I analyzed how human beings can detect them in the absence of any direct instruction and without intention. The foundational work by Katz, Baker, and Macnamara (1974), as well as Gelman and Taylor (1984) showed that children around two years old were able to discern correctly whether an orally conveyed, previously unknown noun referred to an object as a member of a category (common nouns, such as dog) or as an individual entity (proper nouns or proper names, such as Lassie). Toddlers used semantic and syntactic information for their discernment. Semantic information was used to assess the type of thing that was being designated by the noun (a toy animal vs. a cube; a doll vs. a spoon), while syntactic information applied to lexical cues that might be accompanying the noun (Lassie vs. the lassie). This ability to correctly identify spoken proper nouns was verified by testing the nouns against things that the children could not yet name and things that they already could; additionally, the nouns were spoken with and without distracting elements. Jaswal and Markman (2001) confirm that toddlers can learn these nouns with or without ostensible cues (joint attention, looks, direction of the voice, pointing, etc.), both via inference and by being explicitly and unambiguously instructed to do so. Birch and Bloom (2002) proved that children can also identify to what or to whom an anthroponym refers, based on the higher familiarity evidenced by the person using the anthroponym compared to other items.

This early capacity might be part of a cognitive toolset that explains the preliminary results obtained by J. C. Weijer, J. M Weijer, and Yan (2012), who investigated whether names were easier to identify phonetically than common nouns among adults. Their question was based on partial evidence that places both types of words in separate categories: namely, studies showing that the prosody of proper nouns differs phonologically from that of common nouns, at least in
English; the tip-of-the-tongue phenomenon (being aware that a word is known, but being incapable of accessing it) being significantly more prevalent for proper nouns; the different electrophysiological responses elicited by proper nouns and common nouns; the existence of aphasic patients that are selectively unable to produce and understand only proper nouns. Taking the experimental conditions to the extreme, the researchers exposed subjects with very diverse native tongues to full sentences in languages they could not understand (Korean and Swedish), from which they were required to pick out a name. They were also presented with incomplete sentences, which they had to fill out by choosing the best fit between two words. The results suggest that names may present distinctive acoustic traits compared with other words; that the position of names in a sentence interplays with the acoustic traits; that, however, the distinctive phonetical information may not be encoded in the context but in the proper nouns themselves.

With regard to the role of media, Scofield, Williams, and Behrend (2007) is one of the studies to have shown that learning of anthroponyms may start at an early age, even if mediated by technology and without a human presence to facilitate it (e.g., a video with no image of the speaker). Written media is the source of most of the vocabulary a human being acquires when learning to read and write; this process takes place in the absence of a direct instruction and outside of academic environments (Cunningham & Stanovich, 1998). Radio and television are also common sources for anthroponym acquisition (Close, 2004; Greenfield & Beagles-Roos, 1988). Of all the words acquired since birth through the circumstances and ways described, anthroponyms are chronologically among the first (see Hall, 2009, for evidence in English, Mandarin and Cantonese).

Detection, Recording and Retrieval of Morphemes

Previous studies that have investigated the ability of human beings to detect, record and retrieve bounded morphemes (those that are part of a word, such as suffixes) have included anthroponyms among the words analyzed. In studies as early as Berko (1958), 4- and 5-year-old
preschoolers have consistently been shown to be capable not only of detecting without
instruction many of the suffix morphemes found in common English words, but also using them
correctly even in words concocted by the researchers (for instance, they would correctly add the
suffix s to the made-up word wug placed in a nominal function so that it would fit within the
container sentence). Marquis and Shi (2012) and Mintz (2013) showed that infants can detect
suffix morphemes from 11 months and 15 months of age for French and English respectively
(see references for studies involving other languages). Crepaldi, Rastle, and Davis (2010)
showed that adults display great sensitivity about the position of suffix morphemes, and were
able to detect them even in made-up words when these are placed in infrequent positions
relative to the words they learned from the environment. Although the previous observations on
tests based on oral exposition can be applied to tests based on reading and writing methods,
these have accrued specific additional evidence of their own. Rastle, Davis, and New (2004)
found that the segmentation of a word in its morphological components is achieved initially by
processing purely orthographic information, and Lavric, Elchlepp, and Rastle (2012)
corroborated this using brain activation techniques. Quémart, Casalis, and Duncan (2012) found
that French third-graders develop orthographic representations of morphemes and improve their
reading through the use of roots and suffixes, despite the fact that morphological rules are not
explicitly taught until higher grades. The detection and recording of new morphemes in written
form may also be aided by analogical reasoning, through the analysis of morphological links with
previously introduced words (Kirby et al., 2012).

Detection and Recording of Linguistic Regularities and Relationships
The ability of humans to detect, record and retrieve words (including anthroponyms) and
morphemes of different kinds (including those existing in anthroponyms) prompted research into
the cognitive processes that might account for it. This research, in turn, produced a better
understanding of the human ability to extract and to record linguistic regularities and
relationships from the environment: regularities such as the tendency of certain suffixes to
appear in the vicinity of mentions of Russia, and relationships in memories such as the one linking Slavs with chess. For the sake of clarity, regularities and relationships will be treated separately, but there are points in the analysis where their boundaries may be blurry and even arbitrary: when the relationships between linguist elements (e.g., the association of a demonym with a country; the rule whereby a morpheme must be positioned after and not before certain roots) are not learned unambiguously and through direct instruction but through the detection of statistical regularities extracted from the environment, the relationship may be determined by the regularity. This possibility may be illustrated by the following simplified model, borrowing elements from Smith and Yu (2008) and Yu, Smith, Klein, and Shiffrin (2007): let’s suppose the writing of a particular word (for instance, the sequence of letters that spell Russian) is stored in (represented by) a person’s set of neurons, while the stereotypical image of a Russian national is stored in (represented by) a different, disconnected set of neurons. Let’s suppose now that this person is presented with the writing and the image appearing concurrently and repeatedly. Were this co-occurrence to generate a neuronal connection of a certain intensity between both previously unconnected regions, such a connection might represent both a frequency (e.g., through the intensity of the connection if it were proportional to the amount of times both elements co-occurred) and an association (between the written symbols and the referent).

Linguistic Regularities

Words. Saffran, Aslin, and Newport (1996) showed that infants of 8 months of age can detect, record and retrieve words, even from an artificial tongue spoken in a continuous fashion (without pause, accentuation nor prosody), after an exposure of just two minutes. Their conclusions were based on statistical information: the sounds occurring less often (less regular or frequent) tend to signal the boundaries between words, while the sounds occurring more often (more regular or frequent) tend to indicate that the word hasn’t ended yet. Marcus, Vijayan, Bandi Rao, and Vishton (1999) contributed strong evidence of the infants’ use of another cognitive capacity: creating abstract generalizations (algebraic rules) from the detected regularities, an example of
which is the rule *Item I is the same as item J*, where letters represent areas of variable content that may be filled with different data according to the context. Johnson and Jusczyk (2001) found that, even in the face of statistical discrepancy, they can correctly segment words by leveraging other linguistic cues incorporated by the researchers in the speech, such as those reviewed on Gervain and Mehler (2010): stress, syntactic structures, tonal sequences, rythmic variation, legal and illegal phonetic sequences for the studied language, etc.

**Morphemes.** The discovery of regularities is also the main mechanism involved in the detection and recording of morphemes. Marquis and Shi (2012) showed that infants of 11 months of age detected the suffix morphemes in fake words where the variety of the suffixes was lower than that of the roots. The infants detected the morphemes based on the relative frequencies of their co-occurrences, as the lesser amount of suffixes resulted in less variability than roots. It is worth noting that in a previous experiment the fake words were built by combining roots nonexistent in French (for example, *trid-*) with an existent suffix (for example *-e*, which forms the fake word *tride*) or a nonexistent suffix (for example *-u*, which forms the fake word *tridu*): the infants detected the suffix existent in their language, *-e*, but not the inexistent one, *-u* (i.e., they considered *tridu* a morphematically non-divisible word). In another experiment, infants of the same age were exposed to spoken words formed by the fake roots and the fake morpheme (e.g., *tridu*), after a session where they had to listen to other fake words formed by other nonexistent roots and the same fake morpheme (for example, *linchu*, *cradu*, *nadu*, etc.). This familiarization phase was enough to enable them to later detect and record the particle *-u* as a new morpheme, even when they had never heard the root *trid*. Also in this case their detection was based on the linguistic regularities present in the stimuli. Experiments with 8- to 11-year old children based on written media arrived at the same conclusions, but also found that this sensitivity reaches down to the level of the letters of a language, as the children were able to detect and record the probabilities of the letters appearing in a given sequence or order even when the letters didn’t constitute a morpheme (Pacton, Fayol, & Perruchet, 2005).
**Linguistic Relationships**

**Words.** Until not long ago, infants were assumed to depend entirely on the intention and the cues gleaned from the speaker (joint attention, looks, direction of the voice, pointing, etc.), even for the purposes of associating the sound of a word with a referent (what or who is being referred to by the sound). However, Pruden, Hirsh-Pasek, Golinkoff, and Hennon (2006) showed that before reaching their first year of age children can already learn these relationships even implicitly, by looking at individual objects presented on a screen while a recorded voice names them. Smith and Yu (2008) tested this ability on 1-year-old infants under more stringent conditions: the children were presented with more than one object at a time, so that they could not unambiguously link the objects with their referents on the first try. First they were shown a single image with two objects as a voice mentioned two names in succession. Then they were shown an image where one of the previous objects appeared next to a new object, while the voice named them both (precautions were taken to have the position of the new object and the order in which the pair was named change throughout the presentations in a controlled way, so that it was not possible to associate a name with an object by simply noticing regularities such as whether the left object was mentioned first). The infants took less than 4 minutes to learn which object matched which word. The statistical co-occurrence of objects and words is thought to have created a set of strong and weak associations, which grew stronger when certain objects and words appeared together and weaker when they didn’t. After a sufficient number of trials, the strength of the associations would match that of the actual statistic (the number of co-occurrences of each word-object pair would map to the strengthening and weakening of the links between words and objects). The same results have been observed for adults (Kachergis, Yu, & Shiffrin, 2012) and for reading-writing techniques, including meanings (Nagy, Herman, & Anderson, 1985; Frishkoff, Perfetti, & Collins-Thompson, 2010).

**Morphemes.** With regard to the relationship between morphemes and referents or meanings, Berko (1958) determined that preschoolers of 4-5 years of age are also able to unconsciously
discover and assign meanings to morphemes. For example, the suffix \textit{-ing} is always appended to a verbal root with the meaning of “what’s happening presently” (p. 174). The same findings were observed using reading-writing techniques (Pacton et. al, 2005).

\textbf{Associations with Russia}

Why was it conjectured that the Slavic suffixes (not only the Russian ones) would contribute to the significantly higher average success rate of the Russian nationality if, for example, the Russian suffix \textit{-ov} was also disseminated to countries such as Bulgaria, Belarus or Serbia? (As shown in the Appendix B of the main article, Topalov is a Bulgarian surname and Alexandrov is Belarusian). Because there were taken into account environmental characteristics (some of whose effects persist) which might have exposed the subjects to multiple experiences favoring the association of certain suffixes with Russian nationality over other nationalities with surnames including these suffixes.

The group of countries that formed the Union of Soviet Socialist Republics (USSR) for most of the 20th century was also commonly called "Russia" by the major international media outlets (although Russia was just one of the countries in the Union), which led to the adjective "Russian" being used as a synonym of "Soviet" to refer to the USSR (Calvert, 2011; Gerbner, 1991). One reason for this was that the USSR was created after the revolutions that put an end to the Imperial Russian period, which had used the "Russian" name for almost 200 years and which occupied essentially the same territory; another reason was that the assumption of Stalin sparked the so-called \textit{russification} of the USSR, whereby the Russian language was adopted as official and a deliberate Soviet-Russian identification was promoted both within and outside the country (Dietrich, 2005; Mertin, 2008). This identification was disseminated in practically all cultural, sports, political or military activities. In international sports competitions, the sportsmen had to wear distinctive USSR kits and compete under the Union flag, while references to their origins had to be limited to "Soviets" or "Russians". Due to these circumstances, the Slavic
suffixes present in the surnames disseminated by the media were much more likely to appear contiguous or closer to *Russia* than any other nationality, and to the "Russian" demonym than to any other. The probability was much greater if one considers that after the Second World War the United States and the USSR waged an ideological battle through their internal media and their international allies’ media, which led to Russia being mentioned far more frequently than any other Slavic nation (Calvert, 2011; Gerbner, 1991). The studies led by these two authors illustrate the situation: after investigating hundreds of news items devoted to the Summer Olympics of 1952 and 1988 (the first and the last of the Cold War) in large American newspapers, Calvert concluded that «the terms “Soviet” and “Russian” were used interchangeably» (p.92). Gerbner, investigating American television programs and films from the ‘70s and ‘80s (many of which were exported to the rest of the world), found that most of the foreign characters in main roles were Russian, only exceeded in frequency by British and German characters; and that they were "always called Russian, not Soviet" (p. 32). Gerbner clarifies that although research shows that the main influence on children with respect to these concepts has come from the mass media, other disseminators of the same information include parents and the school system.

If that is the case, a player like the Russian Jakovenko (see Table B1 in the Appendix B of the article), would contribute to significantly higher average success rates for the estimation of Russian nationality than for other nations, even though the suffix *enko* is of Ukrainian origin. Otherwise, he would not contribute to these successes (note that the mistakes were not necessarily due to the suffix *enko* being associated with Ukraine in the participants’ memories: it would suffice for the suffix to not be associated with Russia and for Jakovenko to be assigned any other nationality). On the other hand, an examination of Appendix B in the article reveals that the same association with Russia that in general would lead to the significantly higher average successes in the estimation of Russian nationality, would in turn result in more errors in
the estimation of a few players’ nationalities. In particular, Gashimov (Azerbaijan) and Aleksandrov (Belarus) have a typically Russian suffix, but they are not Russian themselves.

**Associations with Chess**

Although chess was considered a popular recreational game during the Russian Empire, with the creation of the USSR it rose to the category of national sport, played by millions of people who participated competitively. Factories and schools had areas destined to its practice, and in many places it was even part of the official education. Thus, the best players were identified precociously, trained in a special way and enrolled to compete in international tournaments (Howard, 1999). The massive practice and media coverage of chess (first within and later outside the Union) was inscribed in the importance the USSR granted to physical education in general as part of an integral education that also aimed to cultivate cognitive, ethical and aesthetic aspects (Mertin, 2008; Zilberman, 1982). Certain sports received special attention with the additional purpose of showing the advantages of socialism over capitalism to the world, so in 1949 it was officially announced the intention to raise the level of all athletes in the USSR, with the aim of achieving short-term world supremacy in most sports (Riordan, 1993). Riordan's study establishes that this policy was decisive in ensuring that "the USSR dominated world chess" (p.248) and that even after its demise in 1991 Russia maintained this supremacy. So:

- Regarding the possibility that the participants may have had in their memories an association between USSR or Russia and good performance in chess: besides the fact that, as mentioned in the article, chess world champions between 1948 and 2007 were almost exclusively Soviet, Russian or from other Slavic nationalities; that a large part of the world's 50 best players between 1970 and 1999 were Soviet, and from 1991 they were from other Slavic countries (Howard, 1999), the USSR also regularly won almost all the Chess Olympiads (which are held every two years and represent the most important team-based international chess tournament) held between 1952 and 1991, while later ones were won mainly by Russia or other Slavic
countries. The news coverage was always wide, especially during the Cold War, since the Chess Olympics were part of the same ideological battle that the United States and the USSR maintained through the mass media in any international sporting event that would pit them against each other, especially on the occasion of the Olympic Games (Calvert, 2011). All this illustrates why the world of chess and its media coverage used Soviet and Russian interchangeably.

- Regarding the possibility that the participants had generated an association between certain suffixes and good chess performance (that is, without having registered any association between Russia or USSR and chess): As mentioned in the article, in addition to the fact that the USSR, Russia and other Slavic countries’ dominance in the chess world spawned countless surnames with Slavic suffixes linked to chess, with some of them (such as the suffix –ov) found in the surnames of several world champions, the world championship matches between Karpov and Kasparov (both with suffix -ov) were the most widely broadcasted chess events of all time, and the reigns of these two champions alone spanned 25 consecutive years. Likewise, Kasparov was (and still remains) in the public eye for his political activism and his books, his public statements about female performance in chess, for continuing to hold the record for the youngest champion in history, for being the first active champion to lose a game (and then a match) against a computer, for having achieved the highest number of ELO points of all time until 2012, etc.; all facts that still make him a frequent object of mass media attention.

**Detection and Processing of Suffixes**

Goldstein and Gigerenzer (1999, p. 39 and Figure 2-1) clarify that they use the term recognition to divide the world into the novel and the previously experienced, and add: "We treat recognition as a binary phenomenon: one either recognizes or does not" (p. 56); "recognition concerns the difference between items in and out of memory" (p.57). They thus define the categories recognized objects and unrecognized objects, further dividing the recognized objects in two: the
merely recognized and the recognized of which something else is known (for example, when deciding which of two recognized cities has more inhabitants, the additional information that one of the cities has universities and the other does not can serve as a cue). Unfortunately, the literature has traditionally assigned different meanings to "cue" without settling on a clear definition. The term can refer to characteristics of an object: the shape of a bird’s head or its color are cues that differentiate it from another bird. It can refer to cognitive functions: cue as a means of accessing a record in memory, or as an identifier in decision-making (see Berretty, Todd, & Martignon, 1999, page 235, for a typical indiscriminate use of cue with both meanings, already in the first two paragraphs of the article). Objects can refer to elements that are presented to a subject (e.g., names of chess players in Campitelli & Labollita, 2010) or to portions of information in memory (examples in Gigerenzer & Todd, 1999). Despite these ambiguities, studies have coincided in that a recognition judgment always refers to the item or object presented, not to one of its characteristics or to memorized portions of information. For example, a subject declares whether they recognize a bird or the name of a city. Further recognition analyses are then performed in relation to the item, not in relation to a characteristic that the subject may have perceived about the item (e.g., the color of the bird) or recovered from his memory in relation to the item, even without perceiving it (e.g., "it has a university"). Studies also agree that subjects must recognize the element or elements presented to them before they can use additional information. The two tasks in Campitelli and Labollita (2010) are part of the enormous set of investigations that have asked participants to judge the recognition of the elements presented, not of any of their characteristics. In a similar vein, linguistic cues (suffixes such as -ov, -ev, -enko) are never disseminated by mediators as isolated elements or objects.

As we have seen, probabilistic environments are systems constituted by entities whose relationship levels are described in this study by their ecological correlation, substitute correlation and validity with the criterion (Gigerenzer & Goldstein, 1996, Gigerenzer & Todd, 1999, Goldstein & Gigerenzer, 2002). However, less information is provided by these
environments about the internal processes of their entities. Just as this study can say little about what happens inside the mediators, not much can be said either about the cognitive processes involved in the detection of suffixes at the perceptual level (when reading surnames) and about their subsequent use (when accessing memory). Not only should additional and more specific tests be carried out; even with them, the field of heuristics currently presents theoretical problems that are difficult to solve when attempting this kind of attributions. Psycholinguistic evidence of basic level, such as the finding that the segmentation of words into their morphological components is primarily achieved by processing purely orthographic information (Rastle, Davis, & New, 2004; ratified later by Lavric, Elchlepp, & Rastle, 2012, using brain activation techniques), does not shed much light upon the area of heuristics related to judgments and decision-making when discussing higher cognitive processes. Therefore, processes such as familiarity, availability and fluency (in addition to recognition), are typical candidates for being considered responsible for the detection and/or the partial or total processing of suffixes. Adding to the difficulty of having to choose among them using the information obtained in this study are the controversies surrounding the definitions, characteristics, limits and superpositions of those processes. Many researchers have defined them using little more than labels; others have presented and investigated them as fully-fledged heuristics; some have defined them conceptually; others through complex mathematical and/or algorithmic models (Gigerenzer & Brighton, 2009; Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1973; see Gigerenzer & Todd, 1999, for a comprehensive review). Even if more or less conventional and simplified meanings of these processes were used, it would be found that there is no agreement about the moment they act (Goldstein & Gigerenzer, 2002; Newell & Fernández, 2006; Pachur & Hertwig, 2006; Schooler & Hertwig, 2005). For example, the model of dual recognition processing, which is divided into a phase of familiarity (a mere sensation of a prior encounter with an object or event) and a phase of proper recognition (the retrieval of the context for that event or element encountered previously), not only continues to incite
controversies as to whether both operations work in parallel or not, but also as to their order of appearance among those who agree that they appear sequentially (see the review by Mandler, 2008). Other difficulties that apply only to recognition can be found in Newell and Fernández (2006), whose authors support the view that recognition is the name given to an attribution that is additional and subsequent to the processing fluidity of a read item (that is, an individual first perceives the level of fluidity with which he processes an element, and only later perceives the element as recognized or not).


