Power law distribution of frequencies of characters in dreams explained by random walk on semantic network

Richard Schweickert[1], Zhuangzhuang Xi[1], Charles Viau-Quesnel[2], and Xiaofang Zheng[1]

[1]Department of Psychological Sciences, Purdue University, USA
[2]Département de psychoéducation, Université du Québec à Trois-Rivières, Canada

Summary. In an individual's dreams some characters occur more frequently than others. In dream series of five individuals, character frequencies follow a power law probability distribution, a distribution often found for contact with people in waking life. Knowing the form of the distribution is important for statistical considerations, because a power law distribution is highly skewed. The form also constrains explanations of how characters are generated in dreams. Character generation is analogous to naming people in a verbal fluency task. We explain the power law with an established model for this task, a random walk on an individual's semantic memory for people and their associations. We demonstrate with simulations that a random walk on such a network can produce a power law character frequency distribution, whether the random walk is self-avoiding or not and whether the network is connected or not.

Keywords: Social network, Zipf's Law, dream characters

1. Introduction

Much is known about the proportion of people in dreams who are family members, strangers, or of other types (e.g., Hall & Van de Castle, 1966; Schredl, 2013). But little is known about a more basic matter. Some people occur more often than others in an individual's dreams. What is the probability distribution of their frequencies?

Naturally occurring quantities often follow a normal distribution, so a distribution with a peak in the middle is plausible. But frequencies of contacts with people in waking life tend to follow a highly skewed distribution with a peak at a tail, known variously as a power law, Zipf, or zeta distribution (Newman, 2005; Johnson, Kemp & Kotz, 2005). A power law distribution for occurrences of people in both waking life and dreams would be consistent with the Continuity Hypothesis, "There is considerable congruence between what a person dreams about at night and what he does or thinks about when he is awake" (Hall & Norbby, 1972, p. 125).

Although correspondence between events in dreams and in waking life is not complete, it is extensive (e.g., Kahan, LaBerge, Levitan & Zimbardo, 1997; Pesant & Zadra, 2005). For discussion see Hobson & Schredl (2011) and commentaries that follow. Clearly some events in dreams have no counterpart in waking life, so, although the Continuity Hypothesis still engenders debate (e.g., Erdelyi, 2017; Domhoff, 2017), research here, like much on the Continuity Hypothesis, is not so much aimed at testing the hypothesis as at learning which aspects of dream and waking life correspond and which do not. Here we are concerned with whether frequency of occurrence of people in dreams has the same form as that often found for frequency of occurrence of people in waking life, a power law; and, if so, how to explain it.

Previous work reports evidence for power laws in dreams. A power law was reported in a talk for frequencies of characters by Schweickert and Xi (2007). Domhoff and Schneider (2008) found a power law for the frequencies of the eight most frequent characters of a middle-aged woman. In dreams of three individuals, Schweickert (2007) found a power law for the number of characters a given character occurred in a dream with. Here we consider a power law for all frequencies of characters of five dreamers.

The Continuity Hypothesis does not specify how a power law would be produced. Here we explain a power law with a random walk model (Han, et al., 2016; Schweickert, et al., 2020). We propose that a memory task, people naming, is a fruitful analog of generation of people in dreams. In the people naming task, a person is simply asked to name people. It is a version of a verbal fluency task, often used as a test of memory impairment. There are several models of the task (Abbott, Austerweil, Griffiths, 2015; Goñi, et al., 2010; Hills, Jones & Todd, 2012; Sung, et al., 2012). We illustrate with simulations how one of the models can account for the frequencies with which people occur in dreams. The model is a random walk on an individual's semantic-memory network for people and their relations (Abbott, Austerweil, Griffiths, 2015; Goñi, et al., 2010). It is natural to consider a random walk because of its affinity to mind wandering (Killeen, 2013), a waking life activity similar to dreaming (Fox, et al., 2013; Domhoff, 2018a, 2018b).

Corresponding address:
Richard Schweickert, Department of Psychological Sciences, Purdue University, 703 Third St., West Lafayette, IN 47907, USA.
Email: schweick@purdue.edu

Submitted for publication: February 2020
Accepted for publication: August 2020
DOI: 10.11588/ijodr.2020.2.71370

2. Method

2.1. Dream Series

Dream reports were those used in a previous study (Han, et al., 2016; see Schweickert, 2019, for data and errata). Reports were acquired from DreamBank.net (Schneider and Domhoff, 2019). Dream series had between 200 and 500 reports (a substantial number but feasible for hand coding), were in English, by adults who were not elderly and had no known psychological disorder. Dreams of a husband and wife pair were excluded. At the time DreamBank was retrieved, five series met these criteria. Alta, female, wrote in the 1980s, early 1990s, and 1997. Arlie, female, wrote in the 1990s. Merri, female, wrote in 1999-2000. Phil, male, wrote in 1971. The Engine Man (called The Natural Scientist on DreamBank) wrote in 1939, with one more report in 1949. Names of dreamers and characters are pseudonyms. More information about the dreamers is available on DreamBank.net, and for the Engine Man in Hobson (1988).

2.2. Dream Coding

Characters in the dream reports were coded with a version of the Hall-Van de Castle (1966) system, slightly modified to be more stringent about when a character is coded as in a dream (Schweickert, 2007). With the Hall-Van de Castle system, a character is coded as in a dream if the character is mentioned in the dream report or if a belonging of the character is present, even if the character is not present in the scene. The system was modified so a character is coded as in a dream if and only if someone in the dream, including the dreamer, had or could have had a social interaction with the character. Characters in a group are ignored, unless the number of individuals in the group is known, or individuals are discussed separately. As in the Hall-Van de Castle system, fanciful person-like entities are considered characters. Animals are not considered characters, but an entity taking a role ordinarily taken by a person, e.g., a fish who is a woman's husband, is considered a character. The dreamer was a character present in every dream. For parsimony the dreamer was not coded as present unless in a metamorphosed form. For the modified system, reliability for coding presence of individual characters as proportion of agreement is .91 (Han, et al., 2016).

3. Results

As an example of characters with different frequencies, here is an excerpt from a dream report of Alta, “Sue is interviewing someone on cable TV.” Sue occurs in three dream reports of Alta. There is no indication that the person denoted “someone” appears in any dream report but this one.

The ten most frequent characters in the dreams of Alta and their frequencies are in Table 1. High frequencies occur rarely; the highest, 22, occurs once. Lower frequencies sometimes occur multiple times. For example, frequency 9 occurs for two characters. The number of times a frequency occurs is its *count*. Table 2 gives the counts for the frequencies of all the characters in Alta's dream reports.

A huge number of characters, 1140, have frequency 1. Such characters often have roles as “extras,” e.g., as passersby, and are common in dreams (e.g., Strauch & Meier, 1996). Empirical data often have power law behavior, but seldom over the entire range of observed frequencies. In particular, there is often a poor fit at low frequencies (Johnson, Kemp & Kotz, 2005; Clauset, Shalizi & Newman, 2009). One reason is that observations are sometimes produced by a mixture of a power law distribution and another distribution. Another reason is that even when data are produced entirely by a power law distribution, counts of low frequencies are variable and difficult to estimate precisely. In practice, because procedures for fitting a power law are sensitive to extreme values, a lower bound x_{min} is usually estimated and a power law is fit to counts greater than it (Clauset, Shalizi & Newman, 2009).

3.1. Fits of Probability Distributions

We began by fitting a power law to all frequencies including frequency 1. Results clearly indicate that frequency 1 is not fit well. We then fit a power law to frequencies except 1 and satisfactory fits were obtained. We also consider an alternative distribution, the Poisson. Here are the details.

3.1.1 Fits Including Frequency 1

We fit a power law in form of the Zipf-Mandelbrot Law (Mandelbrot, 1965) to the count of characters at each frequency including frequency 1. With this distribution, for counts $k = 1, \ldots, N$ the probability of count k is

$$P(k) = 1/c(k + b)^a, \quad (1)$$

where $a > 0$ is the exponent, $b > 0$ is the shift parameter, and c is a normalizing constant, whose value is determined by a, b, and N. If b is 0 and N is infinite, the distribution is sometimes called a Zipf distribution or zeta distribution (Johnson, Kemp & Kotz, 2005).

Parameters were estimated to maximize likelihood. The maximum likelihood estimator of N is the largest observed nonzero count, given for each dreamer in Table 3. Parameters a and b were estimated using Excel Solver. A measure of goodness of fit is G^2 (e.g., Bishop, Holland and Feinberg, 1975). If observations are sampled from the hypothesized distribution, G^2 has approximately a chi square distribution, with an expected value equal to its degrees of freedom. There are three parameters to estimate, a, b, and N, so the degrees of freedom for the N observations are $N - 3$. For each dreamer except Phil, the Zipf-Mandelbrot

*Table 1.*Ten most frequent characters in dreams of Alta

Character	Frequency
My brother	22
Lori	12
My mother	11
My father	9
Bonnie	9
Linda	8
Jenny	8
George	7
My cousin	7
Dan	6

Table 2. Nonzero counts of frequencies in dreams of Alta

Frequency	Count
1	1140
2	29
3	12
4	6
5	4
6	1
7	2
8	2
9	2
11	1
12	1
22	1

law is rejected. (For Alta, $G^2(19) = 63.34$, $p < .001$; for Arlie, $G^2(26) = 74.03$, $p < .001$; for Merri, $G^2(65) = 155.59$, $p < .001$; for The Engine Man, $G^2(9) = 30.19$, $p < .001$; and for Phil, $G^2(45) = 33.92$, n.s.) The bulk of the characters have frequency 1. This bulk is not fit naturally by a power law.

Let's consider an alternative distribution. It is possible that characters appear in dreams so haphazardly that every character appears independently with the same probability. We can test the hypothesis of independence and equal probability because it implies that with a large number of characters, the distribution of character frequencies would be approximately Poisson (e.g., Newman, Strogatz & Watts, 2001). Intuitively, a Poisson distribution might give a satisfactory fit, because it can have both bulk near the lower tail and a long upper tail. It describes counts of frequencies well in some situations (e.g., Bishop, Holland and Feinberg, 1975). When considering a power law, Clauset, Shlizi and Newman (2009) recommend comparing an alternative distribution. In the case of a discrete distribution, as for the frequencies here, they recommend considering the Poisson distribution as an alternative.

In the truncated version of the Poisson distribution, parameters to be estimated are the mean λ and N. The truncated Poisson distribution does not fit well when frequency 1 is included. (For Alta, $G^2(20) = 576.75$, $p < .001$; for Arlie, $G^2(27) = 642.68$, $p < .001$; for Merri, $G^2(66) = 2507.97$, $p < .001$; for The Engine Man, $G^2(10) = 216.93$, $p < .001$; and for Phil, $G^2(46) = 560.27$, $p < .001$). The Poisson distribution is forced to compromise. Its mean λ must be high enough to fit the observed long sparse upper tail, but low enough to fit the dense lower tail. It does not succeed at this. We reject the hypothesis that all characters appear independently with the same probability.

3.1.2 Fits Excluding Frequency 1

Both the power law and Poisson distributions fit badly when frequency 1 is included. Obviously, in the frequency counts in Table 2, the count of frequency 1 is exceptional. Characters appearing in only one dream are usually in roles such as clerk, and seem to be produced by a different probability distribution than characters appearing in more than one dream.

For all dreamers, when frequency 1 is omitted the power law fits well for frequencies 2 to the highest observed frequency, but the Poisson distribution still fits badly, see Table 3 and Figure 1. For the power law G^2 is relatively low for all five dreamers, whereas for the Poisson distribution

Table 3. Fits of Poisson and power laws to frequencies of characters in dream series: Character frequencies greater than 1

		Dreamer				
Distribution		**Alta**	**Arlie**	**Engine Man**	**Merri**	**Phil**
	Dreams	423	208	214	312	220
	Characters	61	20	24	71	41
	N	22	29	12	68	48
Poisson	λ	3.44	7.42	3.65	7.42	4.43
	G^2	40.89	127.14	29.75	809.68	194.46
	df	19	26	9	65	45
	p	< .001	< .001	< .001	< .001	< .001
ZM	a	2.41	1.52	1.85	2.00	2.34
	b	0.23	0	0	0	0
	G^2	11.32	25.97	8.60	62.91	17.50
	df	18	25	8	64	44
	p	n.s.	n.s.	n.s.	n.s	n.s.

Note. For each dreamer, N is the largest observed frequency. For the Poisson distribution, λ is the mean. ZM is the Zipf-Mandelbrot distribution with parameters a and b in Eq. (1). For Alta with b fixed at 0 in Eq. (1), a = 2.28, and G^2 increases to 11.34, df = 19. The slight increase in G^2 is not significant.

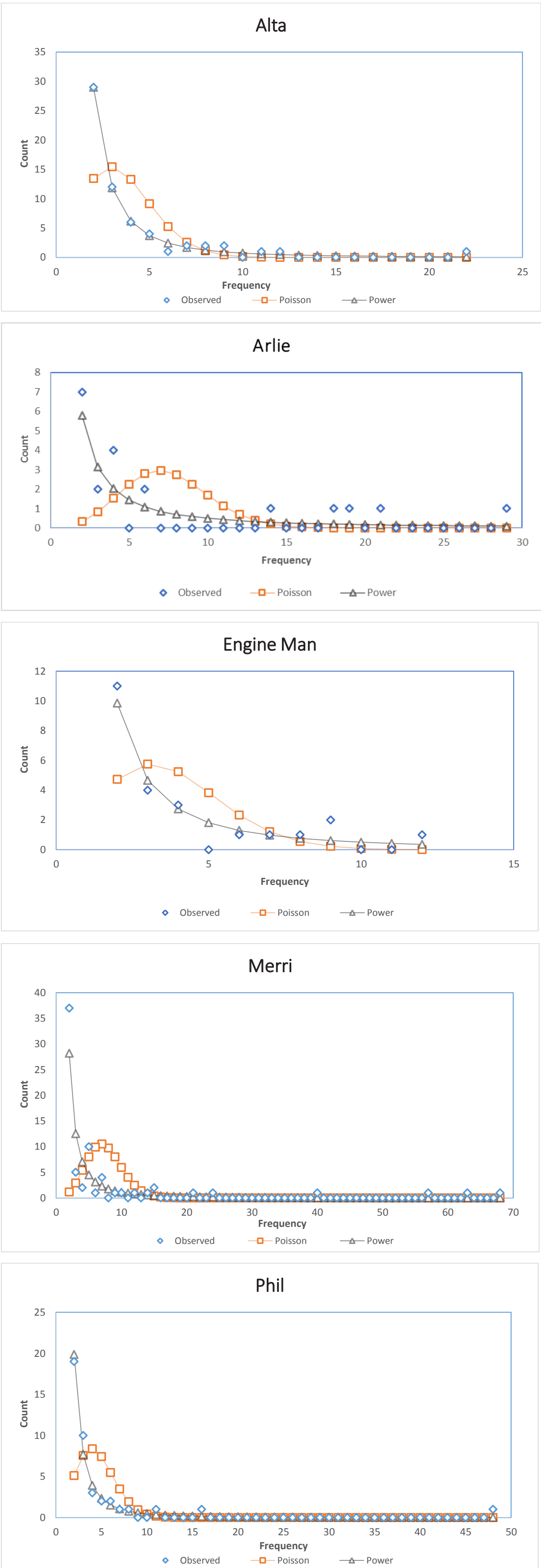

Figure 1. Counts observed and predicted by Poisson and Power Law distributions for five dreamers.

G^2 is relatively high. For frequencies greater than 1 we reject the hypothesis that all characters appear independently with the same probability, but we find the power law fits well.

3.2. Independence of Characters

Here we briefly consider the hypothesis that characters appear independently of one another, regardless of what the distribution of frequencies may be. It is already known that characters do not appear in a dream independently of one another. In an analysis of dream reports from the DreamBank website of a person named Emma, Domhoff (2003, p. 104) found that her husband and her minister appeared together significantly less frequently than predicted if their appearances were stochastically independent. Another example is in dream reports of Merri. The four characters of highest frequency are, in order, the dreamer's sister, Dora; the dreamer's brother, Roger; the dreamer's mother and the dreamer's father. (The pseudonym "Roger" was later changed to "Ruddy" in DreamBank.) Pairs of these characters appear in dreams together more frequently than if independent, except for Dora and Roger for whom independence is not rejected (Schweickert, 2007). Frequencies of characters depend on what other characters are present.

3.2.1 Independence of characters from dream to dream

In addition to considering how characters co-occur in dreams together one can consider how characters occur from one dream to another. We calculated the frequency with which one of the four major characters was followed by one of them in the succeeding dream. Take the case that Mother is followed by Dora, for example. All pairs of dream reports were found such that the character Mother appeared in one dream and Dora appeared in the succeeding dream on the following date. Results for this example are in Table 4.

Occasionally, Merri reports more than one dream from the same night, and we separately calculated the frequency with which one major character followed another in successive dream reports with the same date. For dates with more than one dream report, the first was considered the successor of the last report from the preceding day, and the last was considered the predecessor of the first report from the following day. For example, if there were two dreams on day one in which Mother appeared, and one dream on day two in which Dora appeared, we treated the last from day one and the one report from day two as a successive pair.

Dream reports are numbered in order on the DreamBank website of Schneider and Domhoff, and each has a date. A few discrepancies were found between order and date. Report #24 is two days earlier than #23, #40 is one day earlier

Table 4. Mother and Dora in different successive dreams of Merri

	Dora (sister) Night 2	
Mother Night 1	Present	Absent
Present	10	30
Absent	40	130

Table 5. Fits of power law to frequencies of characters generated by random walks on networks: Character frequencies greater than 2

	Connected		Disconnected	
	Not Self-Avoiding	**Self-Avoiding**	**Not Self-Avoiding**	**Self-Avoiding**
N	82	55	86	44
exponent a	1.45	1.99	1.49	1.59
G^2	96.87	55.43	94.28	46.41
df	78	51	82	40
Significance	n.s.	n.s.	n.s	n.s.

than #39, #44 is one day earlier than #43, and #138 is one day earlier than #137. Dates were assumed to be correct, and the report number order to be incorrect.

There were 42 pairs of successive dreams zero days apart and 210 pairs one day apart. For those one day apart, the largest frequency of successive occurrence is for Dora followed by Dora, for which a permutation T test (done with perm in R) is not significant (p= .58). There is no evidence that appearance of Dora is dependent on her appearance in the preceding dream report. Other pairs of the four major characters occur in succeeding dream reports less frequently than this, and permutation T tests for other pairs are not significant. The result was the same if pairs of dreams from the same date were included. We conclude that appearances of major characters in successive dream reports are independent.

3.3. Model

Generating characters in dreams is similar to generating names of people in the people naming task. People naming and other such verbal fluency tasks are often used for cognitive assessment, because performance declines with dementia (e.g., Henry, Crawford & Phillips, 2004; Rosen, 1980), schizophrenia (e.g., Sung, et al., 2012), and other impairments.

Two main types of models of the verbal fluency task are foraging in a region of a semantic space (e.g., Hills, Jones & Todd, 2012; Rhodes & Turvey, 2007) and random walks on a semantic network (e.g., Goñi, Martincorena, Corominas-Mutra, et al., 2010; Thompson & Kello, 2014; Troyer, Moscovitch & Winocur, 1997). The models make many similar predictions, although differing about details (Abbott, Austerweil & Griffiths, 2015). Here we consider random walks because they are a natural model for mind wandering (e.g., Kennett & Austerweil, 2016), a mental activity akin to dreaming (Fox, Nijeboer, Solomonova, Domhoff & Christoff, 2013; Domhoff, 2018a, 2018b). It is plausible that random walks are carried out on a dreamer's memory network for people because they serve useful functions on online social networks, for example, facilitating later searches (Sarkar & Moore, 2011).

According to a random walk model, a participant in a people naming task has in memory a semantic network of people and associations between them. The people are represented by points, called vertices, and an association between two people is represented by a line, called an edge, between the vertices representing them. In the social network literature, such a network is called the participant's *cognitive social network* (Krackhardt, 1987). Given the instruction to name people, the participant steps from the cue "people" to a vertex in this network, names the person at this vertex, steps at random to an adjacent vertex joined to it by an edge, names the person at that vertex, and so on.

The *degree* of a vertex in a network is the number of vertices adjacent to it. Consider a vertex chosen at random from the entire network, every vertex having the same probability of being chosen. The probability the degree of the chosen vertex is k, for k = 0, 1, . . . is called the *degree distribution* of the network. Friendship networks and other social networks in waking life often have a power law degree distribution (Newman, 2003).

Suppose the dreamer's waking life social network has a power law degree distribution. Suppose the dreamer's cognitive social network is a more or less veridical representation of their waking social network, and in particular has a power law degree distribution. Most people in the cognitive social network are family and friends. But the network also includes fictional characters and celebrities; for example, a dreamer may know that his sister enjoys stories about Hercule Poirot written by Agatha Christie. Finally, suppose during dreaming a random walk is taken through the cognitive social network. When each vertex is reached, the corresponding character appears in the dream.

It is well known that in a random walk on a network, as the number of steps increases the probability a vertex is visited approaches the degree of the vertex divided by the sum of degrees. The assumptions required are mild, such as that the network is finite and connected (e.g., Bonato, 2008, Theorem 5.6). The result is that a random walk can produce a power law distribution of the frequencies with which characters appear in dreams.

With simulations, we consider two situations not meeting the assumptions the result applies to. First, in a random walk, a vertex can be visited more than once, but in a dream, unless it is unusually long, a character does not appear and then reappear. In most dream coding systems a particular character is coded as present only once, even if mentioned more than once in a dream report (e.g., Schredl, 2010). A suitable model for this situation is a *self-avoiding random walk*, in which a vertex is never visited more than once. Second, the result applies to connected networks, but over time some connections in a dreamer's cognitive social network may weaken, so the network becomes disconnected. Our simulations include a disconnected network and self-avoiding random walks.

3.4. Simulations

In the first phase of the simulations, networks with a power law degree distribution were formed, one connected, one disconnected. In the second phase, random walks and self-avoiding random walks were carried out on the networks.

The simulated cognitive social networks formed in the first phase had 150 vertices representing characters, because a rough estimate of the number of people an individual maintains contact with is 150 (Dunbar, 1992). An exponent for the degree distribution of a social network is typically about 2 (Newman, 2003). So a sample of 150 candidate degrees for vertices was taken from a power law distribution with exponent $a = 2$ and smallest possible degree 1. (Parameter b in Eq. (1) was 0.) A connected network with no loop (i.e., edge joining a vertex to itself) or multiple edges between two vertices was formed from the sampled degrees, if possible. If not, samples of candidate degrees were taken until such a network could be formed. Then a disconnected network was formed that had the same degree sequence as the connected network. More details are in the Appendix.

In the second phase, random walks on the networks were carried out in MATLAB. The first vertex was selected with probability proportional to its degree and the corresponding character was recorded as occurring in a dream. In a simulated ordinary random walk, after a vertex was visited the next vertex was selected at random from the vertices adjacent to it, each with the same probability of being selected, and the corresponding character was recorded as occurring in the dream. In a simulated self-avoiding random walk, the next vertex was selected at random from the adjacent unvisited vertices, each with the same probability of being selected; a vertex that was visited was never visited again. In each of the dream series we analyzed, there are many dream reports, with few characters in each. Similarly in the simulations, random walk steps continued until six characters were generated, or, in a self-avoiding random walk, until six characters were generated or no further steps were possible. In each series of random walks, 300 simulated dreams were generated.

If a non self-avoiding random walk on the connected network were infinitely long, the power law (with b in Eq. (1) equal to 0) would fit exactly for all frequencies, e. g., Bonito (2008). With the small number of steps in the simulation of a non self-avoiding random walk on the connected network, the power law does not fit all frequencies well, $G^2 = 181.89$, 81 df, $p < .001$. The small number of observations for characters of low degree leads to imprecise estimates of their counts, sampling error. For example, exactly one character had frequency 1, so the count of frequency 1 is 1, and the power law attempts to fit this. But for frequency 1 the best fitting power law predicts count 25.03 instead of count 1. The power law did not fit all frequencies well for any of the four random walks considered.

As mentioned earlier, in applications the power law usually only fits above a threshold (Clauset, Shalizi and Newman, 2009). Here, for each of the random walks and networks simulated, the power law fits quite well above frequency 2, see Table 5. In each case, 2 is the smallest frequency above which G^2 is not significant. The exponents in Table 5 are reasonable as estimates of the exponent 2 of the degree distribution of the networks that generated the characters, although in our results they underestimate 2.

4. Discussion

The simulations demonstrate that a random walk on a network can produce frequencies of characters that follow a power law distribution above a threshold, in our simulations, above frequency 2. A power law was produced whether the random walk was self-avoiding or not, and whether the network was connected or not.

In the first part of the paper we found that the frequencies of characters in dreams of five individuals follow a power law distribution above a threshold, in our data, above frequency 1. The reason the distribution does not fit the low frequencies differs in the simulations and the data. For the dream data, the power law predicts a lower count than observed, probably because the observed frequencies are produced by a mixture of a power law and another distribution, the latter having a heavy lower tail. For the simulated random walk data, the power law predicts a larger count than observed. With relatively short random walks, low degree characters are sampled rarely, making estimates of their counts imprecise. Nonetheless, in data and in simulations, a power law fits well, except for low frequencies.

The most salient elements in dreams are people (e.g., McNamara, et al., 2005; Tuomin, et al., 2019). Many are not identified or recognized in dream reports. For such people, there is only evidence that they occur in one dream. Those who occur frequently are family, friends and others important to the dreamer. Representations of these people are somehow generated from the dreamer's memory. It is possible that generation is random, in the sense of haphazard or lawless. But this does not happen. Instead, generation is random, but in the sense of following a probability distribution. Our analyses provide evidence that for frequencies above 1, the distribution is a power law, well-known to apply to frequencies of many kinds of events (e.g., Newman, 2005).

A basic aspect of the appearance of people in dreams is their underlying probability distribution, so establishing its form is important in itself. For statistical analyses, it matters that the power law distribution is highly skewed and, if N is infinite, it may not have a finite mean or variance (Johnson, Kemp & Kotz, 2005). Further, the form of the distribution constrains theoretical explanations of how people become generated in dreams from the dreamer's memory. We consider a model in which characters are generated by a random walk on the dreamer's semantic memory of people and their relations, i.e., the dreamer's cognitive social network. Our simulations demonstrate that a random walk on a cognitive social network with a power law degree distribution can generate characters with a power law frequency distribution.

4.1. Mind wandering

A fruitful proposal is that dreaming is a form of mind wandering (Domhoff, 2018a, 2018b; Fox, et al., 2013). Knowledge of mind wandering is growing rapidly, of activities such as spontaneous thought while relaxed, creative thinking, and dreaming (for review, see Smallwood & Schooler, 2015). In addition to empirical work, a few models have appeared, of underlying brain activity in neuroscience (Mittner, et al. 2014), and of switching attention from on-task to mind wandering in cognitive psychology (van Vugt, et al., 2015). But as yet there is little modeling of moment to moment subjective activity. We are not testing the proposal that dreaming is a form of mind wandering, but we are proposing a model inspired by it. A classic notion is that thoughts follows paths of associations of ideas, both when awake (e.g., Hobbes, 1651/1998) and while dreaming (Freud, 1900/1955, p. 284). A random walk on a network is a natural model for this, although there are contending models, foraging (e.g., Hills, Jones & Todd, 2012) and spreading activation (e.g., Collins & Loftus, 1975). It is well established that items can become associated in pairs, so a network of associations is implied. This associative network is part of an individual's larger semantic memory network, which includes knowledge of people and the world. There is evidence that some tasks relying on semantic memory are conducted by random walks on networks; for example, in semantic fluency tasks (Griffiths, Styvers & Firl, 2007; Abbott, Austerweil, Griffiths, 2015; Goñi, et al., 2010) and in creative thinking (Kenett & Austerweil, 2016). Mildner and Tamir (2019) proposed a random walk as a model of mind wandering while awake. Evidence from mind wandering activities in waking life encourages pursuit of random walk models of mind wandering in dreaming.

Here we showed that a random walk on a cognitive social network is able to generate characters having a power law frequency distribution, found empirically in dreams of individuals. A random walk on a semantic network also is natural for modeling other aspects of dreams. Hobson, et al. (2000) note that in the brain during REM sleep the dorsolateral prefrontal cortex is less active than while awake. One consequence is limited memory during dreaming. A random walk represents this naturally with the memoryless property, the assumption that the next step depends on the current vertex, but not on previously visited vertices. Another consequence, Hobson, et al. (2000) say, is that complete scene shifts occur unnoticed in dreams. The sudden shifts can be modeled by a switch random walk; at any step there is a probability of jumping to a random vertex in the network rather than to an adjacent vertex. Such random walks are models of verbal fluency tasks (e.g., Goñi, et al., 2010; Griffiths, et al., 2007). Finally, following remote associations in dreams is naturally modeled by a biased random walk, in which probabilities of transitions are higher to adjacent vertices that have fewer neighbors (Han, et al. 2016).

4.2. Future work

To test random walks in future work, more content details will be needed. A typical verbal fluency task lasts one minute, with implicit or explicit instruction to not repeat an item. Longer trials and more of them would be useful, and repeated items would allow testing a power law for frequencies of emitted items. The same is true for free association tasks, in which a participant is asked to emit words that come to mind in response to a cue word (e.g., Kuška, Trnka, Kuběna & Růžička, 2016). Also useful would be more reports of subjective experience during mind wandering when awake; see review by Antrobus (2018). When daydreaming, people often generate scenarios with people. If generation is similar to that in dreams, the frequencies of generating various people would follow a power law.

Patients are sometimes observed in a hospital for several days with electrodes temporarily implanted in the brain to locate epileptic foci. While the patient is awake, some electrodes in or near single neurons in the hippocampus, entorhinal cortex and amygdala are found to respond preferentially to visual images of particular people (Kreiman, Koch and Fried, 2000). One hypothesis suggested by work here is that if recordings were made from these electrodes while the patient is awake or in REM sleep, the frequencies with which the various cells are active would follow a power law. (By frequency, we do not mean firing rate, we mean how often the cell is active per unit time in REM sleep.) It would be intriguing if at about the time certain neurons were active their corresponding characters appeared in dream reports. Celebrities and cartoon characters sometimes appear in dreams, and it is worth noting that the preferred images for some neurons are celebrities and cartoon characters.

Dream reports have some advantages over wake reports. On the one hand, a person mind wandering while awake can give an ongoing verbal report, not possible while dreaming. On the other hand, after 15 minutes or so of uninterrupted mind wandering awake a person sometimes has difficulty remembering content in detail, but after a similar interval of dreaming dramatic images often linger, affording report of detail, albeit imperfect. For investigators of mind wandering, dream reports are a rich source of detail.

4.3. Limitations

Coding of characters is time consuming, so our data are from a small set of five adult dreamers. Results may not generalize to other people, in particular to those whose dream series we did not consider because they were children, elderly or have a psychological disorder.

Random walk simulations were conducted on one connected network and one disconnected network. They are sufficient for demonstrating that random walks can produce power law character frequency distributions, but results may not generalize to networks not considered.

Although the model can account for the frequency distribution of characters in dreams, it remains to be seen whether it can also account for other details, such as the small world structure of networks formed by linking two characters if they are in a dream together (Schweickert, 2007). Settings, emotions and other elements besides people are generated in dreams but not considered in the model. Asking dreamers about sources of elements in their dreams reveals waking life sources with patterns such as networks with small world structure (Barcaro & Carbencini, 2018) and systematic delays between occurrences of sources in waking life and their occurrences in dreams (e.g., Eichenlaub, et al., (2018). Moreover, events in dreams form dramas with a story-like organization that becomes more complex later in the night (Cipolli, et al, 2015). A random walk does not by itself produce these patterns.

5. Summary and Conclusions

In dream series of five individuals, character frequencies follow a probability distribution often found for contacts with people in waking life, a power law. A power law restricts explanations of how characters are generated in dreams. Generating characters in dreams is analogous to naming people in a verbal fluency task. A model for people naming is a random walk on an individual's semantic memory for people and their relations. Our simulations demonstrate that a random walk on such a network with a power law degree distribution can generate characters with a power law frequency distribution, whether the network is connected or not, and whether the random walk is self-avoiding or not. We note that in the simulations and the data a power law does not fit well at very low frequencies, an exception that often occurs in applications.

Dreams are remarked on more often for their peculiarities than their regularities, but there are many regularities to account for. Here we propose an account of one of them, the lawfulness of character frequencies.

Acknoledgements

We thank Hye Joo Han for coding and useful discussions and G. William Domhoff for much helpful advice.

References

Abbott, J., Austerweil, J., Griffiths, T. (2015). Random walks on semantic networks can resemble optimal foraging. Psychological Review, 122(3), 558–569. DOI: 10.1037/a0038693

Antrobus, J. S. (2018). How does the waking and sleeping brain produce spontaneous thought and imagery, and why? In K. C. R. Fox & K. Christoff (Eds.), The Oxford handbook of spontaneous thought (pp. 123-132). Oxford: Oxford University Press.

Barcaro, U., & Carbencini, C. (2018). Network properties of dream sources. International Journal of Dream Research, 11, 120-126. DOI: 10.11588/ijodr.2018.2.42736

Bishop, Y. M. M., Feinberg, S. E., & Holland, P. W. (1975). Discrete multivariate analysis. Cambridge, MA: MIT Press.

Bonato, A. (2008). A course on the web graph. Providence, RI: American Mathematical Society.

Cipolli, C., Guazzelli, M., Bellucci, C., Mazzetti, M., Palagini., L., Rosenlicht, N, & Feinberg, I. (2015). Time-of-night variations in the story-like organization of dream experience developed during rapid eye movement sleep. Journal of Sleep Research, 24, 234-240. https://doi.org/10.1111/jsr.12251

Clauset, A., Shlizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. SIAM Review, 51, 661-703.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. Psychological Review, 82, 407-428.

Domhoff, G. W. (2017). The invasion of the concept snatchers: The origins, distortions, and future of the continuity hypothesis. Dreaming, 27(1), 14-39.

Domhoff, G. W. (2018a). Dreaming is an intensified form of mind-wandering, based in an augmented portion of the default network. In K. C. R. Fox & K. Christoff (Eds.), The Oxford handbook of spontaneous thought (pp. 355-370). Oxford: Oxford University Press.

Domhoff, G. W. (2018b). The emergence of dreaming: Mind wandering, embodied simulation, and the default network. New York: Oxford University Press. Domhoff, G. W., & Schneider, A. (2008). Studying dream content using the archive and search engine on DreamBank.net. Consciousness and Cognition, 17, 1238-1247.

Domhoff, G. W., & Schneider, A. (2008). Studying dream content using the archive and search engine on DreamBank.net. Consciousness and Cognition, 17, 1238-1247 https://doi.org/10.1016/j.concog.2008.06.010

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. Journal of Human Evolution, 22, 469-493. DOI: 10.1016/0047-2484(92)90081-J

Eichenlaub, J. B., van Rijn, E., Phelan, M., Ryder, L., Gaskell, M. G., Lewis, P. A., Walker, M. P., & Blagrove, M. (2019). The nature of delayed dream incorporation ('dream-lag effect'): Personally significant events persist, but not major daily activities or concerns. Journal of Sleep Research, 28, e12697, DOI: 10.1111/jsr.12697

Erdelyi, M. H. (2017). The continuity hypothesis. Dreaming, 27(4), 334-344.

Fox, K. C. R., Nijeboer, S., Solomonova, E., Domhoff, G. W., & Khristoff, K. (2013). Dreaming as mind wandering: evidence from functional neuroimaging and first-person content reports. Frontiers in Human Neuroscience, 7, 412, doi.org/10.3389/fnhum.2013.00412

Freud, S. (1955). The interpretation of dreams. (J. Strachey, Trans. Ed.) NY: Basic Books, Inc. (Original work published 1900)

Goñi, J., Martincorena, I., Corominas-Murtra, B., Arrondo, G., Ardanza-Trevijano, S., Villoslada, P. (2010). Switcher-randomwalks: a cognitive-inspired mechanism for network exploration. International Journal of Bifurcation and Chaos, 20(03), 913–922. https://doi.org/10.1142/S0218127410026204

Griffiths, T. L., Styvers, M., Firl, A. (2007). Google and the mind: Predicting fluency with PageRank. Psychological Science, 18(12), 1069–1076

Hall, C. S., & Van de Castle, R. L. (1966). The content analysis of dreams. NY: Appleton, Century, Crofts.

Hall, C. S., & Nordby, V. J. (1972). The individual and his dreams. NY: Signet.

Han, H. J., Schweickert, R., Xi, Z., & Viau-Quesnel, C. (2016). The cognitive social network in dreams: Transitivity, assortativity, and giant component proportion are monotonic. Cognitive Science, 40, 671-696. DOI: 10.1111/cogs.12244

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. Neuropsychologia, 42(9), 1212-1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Hills, T.T., Jones, M.N., Todd, P.M. (2012). Optimal foraging in semantic memory. Psychological Review, 119(2), 431–440. DOI: 10.1037/a0027373

Hobbes, T. (1998). Leviathan. Oxford, England: Oxford University Press. (Original work published 1651).

Hobson, J. A. (1988). The dreaming brain. New York: Basic Books.

Hobson, J. A., Pace-Schott, E. F., & Stickgold, R. (2000). The dreaming brain: Toward a cognive neuroscience of conscious states. Behavioral and Brain Sciences, 23, 793–1121.

Johnson, N. L., Kemp, A. W., & Kotz, S. (2005). Univariate discrete distributions (3rd ed.). Hoboken: Wiley Interscience.

Kahan, T. L., LaBerge, S., Levitan, L., & Zimbardo, P. (1997). Similarities and differences between dreaming and waking cognition: An exploratory study. Consciousness and Cognition, 6(1), 132-147.

Kenett, Y. N., & Austerweil, J. L. (2016). Examining Search in Low and High Creative Individuals with Random Walks. In A. Papafragou, et al. (eds.). Proceedings of the 38th Annual Meeting of the Cognitive Science Society, pp. 313–318, Cognitive Science Society, Austin, TX.

Killeen, P. R. (2013). Absent without leave; A neuroenergenic theory of mind wandering. Frontiers in Psychology: Personality and Social Psychology, https://doi.org/10.3389/fpsyg.2013.00373

Krackhardt, D. (1987). Cognitive social structures. Social networks, 9(2), 109–134. doi:10.1016/0378-8733-(87)90009-8.

Kreiman, G., Koch, C., Freid, I., 2000. Category-specific visual responses of single neurons in the human medial temporal lobe. Nature Neuroscience, 3, 946-953.

Johnson, N. L., Kemp, A. W., & Kotz, S. (2005). Univariate discrete distributions (3rd ed.). Hoboken: Wiley Interscience.
Kuška, M., Trnka, R, Kuběna, A. A., Růžička, J. (2016). Free Associations Mirroring Self- and World-Related Concepts: Implications for Personal Construct Theory, Psycholinguistics and Philosophical Psychology. Frontiers in Psychology. 7, Article 981. doi:10.3389/fpsyg.2016.00981. PMC 4928535. PMID 27445940
Mandelbrot, B. (1965). "Information Theory and Psycholinguistics". In B.B. Wolman and E. Nagel (ed.). Scientific psychology. NY: Basic Books.
Mittner, M., Boekel, W., Tucker, B. M., & Heathcote, A. (2014). When the brain takes a break: A model-based analysis of mind wandering. The Journal of Neuroscience, 34, 16286-16295.
McNamara, P., McLaren, D., Smith, D., Brown, A., & Stickgold, R. (2005). A "Jekyll and Hyde" within: Aggressive versus friendly interactions in REM and non-REM dreams. Psychological Science, 16, 130-136. https://doi.org/10.1111/j.0956-7976.2005.00793.x
Newman, M. E. J. (2003). The structure and function of complex networks. SIAM Review, 45, 167-256. https://doi.org/10.1137/S003614450342480
Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. Contemporary Physics, 46, 323-351, DOI: 10.1080/00107510500052444
Newman, M. E. J., Strogatz, S., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. Physical Review E, 64, 026118.
Pesant, N., & Zadra, A. (2005). Dream content and psychological well-being: A longitudinal study of the continuity hypothesis. Journal of Clinical Psychology, 62(1), 111-121.
Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. Physica A, 385, 255-260. https://doi.org/10.1016/j.physa.2007.07.001
Rosen, W. G. (1980). Verbal fluency in aging and dementia. Journal of Clinical Neuropsychology, 2, 135-146. https://doi.org/10.1080/01688638008403788
Sarkar, P., & Moore, A. W. (2011). Random walks in social networks and their applications: A survey. In C. C. Aggarwal (ed.), Social network data analytics (pp. 43-78). NY: Springer. DOI 10.1007/978-1-4419-8462-3_3
Schneider, A., & Domhoff, G. W. (2019). DreamBank. Available at: www.dreambank.net. Accessed June 20, 2005.
Schredl, M. (2010). Dream content analysis: Basic principles. International Journal of Dream Research, 3, 65-73. https://doi.org/10.11588/ijodr.2010.1.474
Schredl, M. (2013). Dreams of core family members in a long dream series. International Journal of Dream Research, 6, 114-118. https://doi.org/10.11588/ijodr.2013.2.11055
Schweickert, R. (2007). Properties of the organization of memory for people: Evidence from dream reports. Psychonomic Bulletin & Review, 14, 270-276. https://doi.org/10.3758/BF03194063
Schweickert, R. (2019). Data for The Cognitive Social Network in Dreams. Retrieved from osf.io/69msb.
Schweickert, R., Han, H., Xi, Z., & Viau_Quesnel, C. (2020) Dreaming as a Random Walk through the Dreamer's Associative Network. In M. Vitevitch (Ed.), Network science in cognitive psychology (pp. 80-96). NY:Routledge.
Schweickert, R., & Xi, J. (2007, June). Frequency distributions of characters in dreams. International Association for the Study of Dreams. Paper presented at the meeting of the International Association for the Study of Dreams, Sonoma State University, Sonoma, CA, June.
Smallwood & Schooler (2015). The science of mind wandering: empirically navigating the stream of consciousness. Annual review of psychology, 487-518. doi:10.1146/annurev-psych-010814-015331
Strauch, I., & Meier, B. (1996). In search of dreams: Results of experimental dream research. Albany: State University of New York Press.
Sung, K., Gordon, B., Vannorsdall, T. D., Ledoux, K., Pickett, E. J., Pearlson, G. D., & Schretlen, D. J. (2012). Semantic Clustering of Category Fluency in Schizophrenia Examined with Singular Value Decomposition. Journal of the International Neuropsychological Society, 18, 565–575. doi:10.1017/S1355617712000136
Thompson, G. W., & Kello, C. T. (2014). Walking across Wikipedia: a scale-free network model of semantic memory retrieval. Frontiers in Psychology, 5, Article 96, 1-9. doi: 10.3389/fpsyg.2014.00086
Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology, 11, 138-146. https://doi.org/10.1037/0894-4105.11.1.138
Tuominen, J., Stenberg, T., Revonsuo, T., & Valli, K. (2019). Social contents in dreams: An empirical test of the Social Simulation Theory. Consciousness and Cognition, 69, 133-145. https://doi.org/10.1016/j.concog.2019.01.017
van Vugt, M., Taatgen, N., Sackur, J., & Bastian, M. (2015). Modeling mind-wandering: a tool to better understand distraction. In N. Taatgen, M. van Vugt, J. Borst, & K. Mehlhorn (Eds.), Proceedings of the 13th International Conference on Cognitive Modeling. (pp. 252). Groningen: University of Groningen.

Appendix

Simulations

A random sample with replacement of 150 candidate degrees was generated from a power law distribution with exponent a = 2 and shift parameter b = 0, using command rzipfman in R package tolerance. Not every sample of candidate degrees is graphical, that is, possible as the degree sequence of a network, because the sum of degrees in the degree sequence of a network must be even. The sample was tested with the command is_graphical in R package igraph. If it failed, a new sample was generated. This continued until a graphical degree sequence was obtained. Not every graphical degree sequence is the degree sequence of a simple connected network, i. e., a network with no loops or multiple edges. If possible for the obtained degree sequence, a simple connected network was obtained from it with command sample_degseq in R package igraph. If a simple connected network was not possible for the degree sequence, a new random sample of candidate degrees was taken. The procedure was followed until a random connected network was obtained. The network obtained with the command is sampled uniformly from the possible simple connected networks with the degree sequence.

From the same degree sequence a disconnected network was obtained with command sample_degseq. We note that the disconnected network produced by this command is not chosen uniformly at random from the possible ones.

.