Read Writ (2019) 32:13–31
https://doi.org/10.1007/s11145-017-9727-3

The impact of image characteristics on written naming in adults

Patrick Bonin[1] · Alain Méot[2] · Betty Laroche[3] · Aurélia Bugaiska[1] · Cyril Perret[3]

Published online: 8 February 2017

Abstract The present study was aimed at investigating whether and how image characteristics influence written naming performance in adults. In three different sessions, participants had to quickly write down the names of pictured objects on a graphic tablet. Across sessions, the picture format was different, but the to-be-named objects were the same: There were black-and-white pictures (Snodgrass & Vanderwart's [SV] 1980 drawings), grayscale and colored pictures of the SV drawings as provided by Rossion and Pourtois (2004). Linear-mixed models (LMM) were used to analyze written latencies. The main findings were the following: (1) Colorized pictures yielded shorter written naming latencies than line drawings with the grayscale pictures being situated between the two; (2) Both within- and between-picture format LMM revealed reliable effects of name agreement, objective word frequency, frequency trajectory (the effect was marginal in the grayscale condition), and imageability on written latencies. The influence of image agreement was, however, less stable (reliable only in the colorized condition in the within-picture format LMM analysis; significant with both line drawings and their colorized version only in the between-picture format LMM analysis); (3) None of the interactions with picture format reached significance except the interaction of Image agreement with Picture format. In line with Bonin, Roux, Barry, & Canell (2012b), the findings support a limited-cascading account of written word production.

✉ Patrick Bonin
Patrick.Bonin@u-bourgogne.fr

[1] LEAD-CNRS UMR 5022, Pôle AAFE, Université de Bourgogne Franche-Comté, Esplanade Erasme, 21065 Dijon Cedex, France

[2] Université Clermont Auvergne, CNRS, LAPSCO, 63000 Clermont-Ferrand, France

[3] CeRCA-CNRS UMR 7295, Université de Poitiers, Poitiers, France

Keywords Image characteristics · Written naming · Determinants of picture naming

Introduction

In the present paper, we addressed a simple question: Do surface image characteristics influence written naming performance in adults? More particularly and simply stated, our aim was to determine whether the use of black-and-white drawings, gray-level texture drawings or colored lined drawings for the same set of to be-named objects makes a difference to the temporal dynamics of written naming and to naming accuracy.

Producing the name of a pictured object involves several processing stages which consist in different distinct processes and representations. First, there is an initial level of perceptual (visual) identification. The physical characteristics of the picture, such as shape, color or surface details, are encoded. What is represented by the picture then has to be recognized. At this stage, the target concept is activated along with related semantic information, i.e., concepts that are semantically related to the target concept are activated (e.g., *cat*, *wolf*, *fox* for the target *dog*). Semantic representations also entail the activation of stored knowledge corresponding to functional, associative and other properties of the to-be-named object. From the conceptual level, activation spreads to several lexical entries, often referred to as lemmas (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). The target orthographic word-form then gets activated along with individual graphemes that are used by several peripheral processes to produce, for instance, a visible written trace on a sheet of paper.

Picture naming is an experimental technique that has often been used to investigate the processes and the representations that are involved in conceptually-driven naming (Bonin, Roux, & Barry, 2012a). Of course, the use of this technique requires that researchers have at their disposal a large number of different pictures. Different databases corresponding to pictures of objects (or of actions) have been designed and norms have been collected from these pictures. The norms refer either to the picture characteristics (e.g., visual complexity) or to their names (e.g., lexical frequency or age of acquisition) and they have been collected in different languages and cultures (e.g., Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Snodgrass & Vanderwart, 1980; Tsaparina, Bonin, & Méot, 2011). Different picture databases are now available to researchers and they represent indispensable tools for selecting materials for the design of picture naming studies. The most often used picture stimuli are the Snodgrass and Vanderwart (1980) pictures (SV), which consist of 260 line drawings of common objects. Rossion and Pourtois (2004) transformed the original SV drawings in order to provide colored and grayscaled versions of them. More recently, photographs of objects have been provided accompanied by the norms collected for them (e.g., Brodeur et al., 2012; Brodeur, Guérard, & Bouras, 2014). Likewise pictures of objects that are richer in terms of visual details (e.g., texture, color) are now available. Thanks to the normative information that has been collected for pictures and their names, we know better what the main determinants of naming speed and accuracy in both spoken and written naming are (Alario et al., 2004; Bonin, Chalard, Méot, & Fayol, 2002). The general idea has been to try to link certain determinants of naming speed (and to a lesser extent accuracy) to certain processes and representations involved in spoken and written naming, respectively. Likewise, in written naming, which is the focus of the present study, Bonin, Chalard, Méot, & Fayol (2002) have examined the influence of a set of predictors pertaining either to the visual format of the pictures (visual complexity), to the matching of the output of the visual analysis with stored structural representations (image agreement), to the familiarity of the concepts (conceptual familiarity) or to their semantic richness (imageability), to the matching of the concept with the name used to refer to it (name agreement), or finally, to the picture names themselves (objective word frequency, age-of-acquisition (AoA), or length). They found that name agreement and AoA were the most important determinants of naming speed (see also Alario et al., 2004). The observation that both spoken and written preparation latencies are influenced by the same set of predictors has been taken to suggest that similar processes and representations are involved in both output modalities, an assumption that has received further confirmation through the use of EEG data (Perret & Laganaro, 2012).

To our knowledge, no study to date has performed an in-depth comparison of whether and how pictures having different visual characteristics/surface details influence the involvement of the processes and representations underpinning written naming. This is an important theoretical issue that we addressed in the present study. Indeed, pictures with different formats (e.g., black-and-white drawings, colored drawings, photographs) vary as a function of the richness of their surface detail (they contain more or less details, colors, textures; they are more or less visually complex [see below for rated visual complexity scores concerning the three picture formats used here]). The richness of the perceptual input provided by the to-be-named pictures may possibly alter the temporal dynamics of the written naming process as well as the accuracy of the written outputs. In the field of word reading, an increasing number of studies have attempted to investigate the influence of semantic richness on visual word recognition. Indeed, there are many ways to be rich (Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008), but in the broad sense, richness in the context of words refers to words that provide more information at the semantic/conceptual level. Likewise, a number of semantic variables (e.g., imageability, number of semantic features) have been examined in several lexical processing tasks (e.g., lexical decision, word reading). It has generally been found that words that are semantically rich (e.g., in terms of the number of semantic features or verbal associates) are processed faster than less semantically rich words (Pexman et al., 2008). By analogy with (semantic) richness effects in visual word recognition, we may anticipate that, in written naming, when this relates to the same concepts as those depicted by the pictures, pictured objects that are richer in terms of surface detail should be processed faster and more accurately than objects that are perceptually simpler. Indeed, some studies are able to provide some insights on this issue. Several studies have investigated the impact of surface detail on recognition and/or naming performance (e.g., Bramão, Faísca, Petersson, & Reis, 2010; Brodie, Wallace, & Sharrat, 1991; Heuer, 2016; Rossion & Pourtois, 2004). However and importantly, they were not concerned with the issue of exactly how the perceptual characteristics of pictures influence the dynamics of the written naming processes (see below), which is the focus of the present study. For instance, Rossion and Pourtois (2004) reported that the addition of color to the SV pictures improved both the accuracy and the speed of spoken naming, and this was the case not only for objects with diagnostic colors, such as a lemon or a banana, or structurally similar objects, such as fruits or vegetables, but was also observed for artifacts of all types. The inclusion of texture and shading in the SV drawings only slightly improved the naming accuracy as assessed by name agreement. Finally, it has been shown that when objects are depicted in incongruent colors they are named more slowly than when they are depicted with more protypical colors (Therriault, Yaxley, & Zwaan, 2009).

A key theoretical issue is whether surface stimulus characteristics influence only prelinguistic stages involved in object naming (perceptual identification and semantic acess) as predicted by serial-discrete models (Levelt et al., 1999) or limited-cascading models (Bonin, Roux, Barry, & Canell, 2012b), or whether their influence extends beyond these prelinguistic processes and exerts an impact on linguistic processes, as predicted by full-cascading models (Humphreys, Riddoch, & Quinlan, 1988). Stated differently, the theoretical issue is whether the richness of the perceptual input influences only prelinguistic stages of written naming or cascades to linguistic stages. In the present study, the investigation of this issue was made possible using linear-mixed models (LMM). In a previous study, Bonin, Roux, Barry, & Canell (2012b) explored whether difficulties at the perceptual level had an influence on lexical stages of written naming as predicted by a full-cascaded view of word production. In this study, the variables used by Bonin et al. (2012b) were chosen for their reliability in indexing specific processing levels in written word production. Following the logic of Humphreys et al. (1988) (see also Griffin & Bock, 1998), the authors tested whether their effects were interactive or additive. Interestingly, in one study (Study 1), they presented pictures of objects in color or in black and white and the object names were early-acquired or late-acquired (as indexed by frequency trajectory measures, see below). The rationale was that if activation flows in a cascading manner between two particular processing levels, then the effects of the variables should interact. In contrast, if the transmission of information between the levels is serial and discrete, then the effects of the two variables should be additive. The authors found that the preparation latencies were faster for colorized than for black-and-white drawings, and also faster for early-acquired object names than for late-acquired names. However, the two variables of picture format and 'AoA' did not interact, suggesting that activation spreads serially from the perceptual and conceptual levels to the word-form level in written word production. However, Bonin et al.'s (2012b) study is limited by the small number of items that were used. Also, given that as they adopted a factorial approach in which a large number of variables were controlled for, the excessive matching of other variables may have resulted in the selection of unusual stimuli. Moreover, only a small number of interactions between the picture characteristic variable and certain lexical variables were examined. Thus, the issue of whether the influence of visual surface characteristics acting at the perceptual identification level cascades to the linguistic levels involved in written naming requires a more thorough investigation.

In the study reported below, the same participants were involved in three different writing sessions that took place within one month and that were separated by at least one week. They had to produce the names of objects corresponding to pictures by writing them down as fast as possible while remaining accurate. In each writing session, the same set of object names had to be produced (in a different random order). However, the pictures took the form of either black-and-white drawings (SV pictures), grayscale drawings or the colorized version of the SV drawings as provided by Rossion and Pourtois (2004). The strength of our study lies in the fact that the same participants named the same concepts from different picture formats. Such a design makes it possible to test interactions between this variable of interest and a set of different important psycholinguistic variables. We are not aware of any study in written naming that has adopted this approach. Finally, in addition to providing important findings at a theoretical level, from a methodological point of view, the results should help researchers choose the picture type that is the most appropriate to their research purposes.

Method

Participants

Twenty-nine undergraduate students (3 male students; mean age, 20.30 years; range 18–41 years) of psychology at the University of Bourgogne (Dijon, France) participated in this experiment and were given course credits. All were right-handed French native speakers, without any visual, motoric or language impairment.

Stimuli

We initially selected 179 line drawings from the Snodgrass and Vanderwart (SV) (1980) database. The norms corresponding to the SV pictures and their names were taken from Alario and Ferrand (1999). The characteristics of the grayscale and colored pictures were taken from Rossion and Pourtois (2004). However, we decided to collect new name agreement scores for both the grayscale pictures and the colorized pictures because we found that when Rossion and Pourtois' (2004) name agreement scores were used, they turned out to be uncorrelated with naming latencies taken from the Bonin et al. (2002) study. To collect name agreement, the participants (there were 61 first-year psychology students [N = 31 for the grayscale pictures and N = 30 for the colored pictures] who were taken from the same pool as those involved in the main experiment but all were different) were told that they would see a picture and that they had to write down the first name (which could sometimes consist of more than one word) that came to mind on the answer sheet. When they could not provide the name of the picture, they were asked to indicate the reason (e.g., they did not recognize the object). The number of alternative names provided for a particular picture across participants was recorded and used to compute name agreement scores. In order to avoid having a high number of naming errors, all the pictures retained for inclusion in the naming study had a name agreement score higher than 66.6%. The pictures were displayed in a 10 $\times$ 10 cm square on a computer screen. The values for objective word frequency, number of letters and for orthographic neighborhood (as provided by Orthographic Levenshtein Distance 20: Old20) were taken from the LEXIQUE database (New, Pallier, Brysbaert, & Ferrand, 2004). Phonology-to-Orthography (PO) consistency scores and bigram frequency were taken from the MANULEX INFRA database (Peereman, Lété, & Sprenger-Charolles, 2007). Childhood word frequency values, which are used to compute the frequency trajectory and the cumulative frequency of the words (see below), were taken from the MANULEX database (Lété, Sprenger-Charolles, & Colé, 2004).

Apparatus

The experiment was run using Experiment Builder V 1.1. on Dell Pro (with a 200 Hz screen). The computer controlled the presentation of pictures and recorded the latencies. The participants sat in front of a digitizer (Wacom Intuos 3, sampling frequency 200 Hz, accuracy 0.02 mm) that was connected to the laptop. A sheet of paper with 23 lines placed above the graphic tablet and an Intuos inking contact pen was used to record the graphic latencies.

Procedure

The procedure consisted of three picture naming sessions separated by at least one week. For any given participant, a naming session lasted about 40 min and was run with a specific picture format (colorized, grayscale or black-and-white drawings). The 179 pictures were presented randomly with a short break provided every twenty-three trials. The breaks also allowed us to change the sheet of paper on the graphic tablet. The order of administration of the three picture formats across participants was defined by random permutations of the tasks (there were 3! possible orders and five participants per order except for one order that was tested on four participants), with one format being administrated per session.

The participants carried out the experiment individually in a soundproof room. They sat in front of the screen at a distance of about 70 cm. The instructions were given orally to the participants. The participants were told to write down the name of each picture as quickly and as accurately as possible. Moreover, they were told to position the pen above the start of the line and not to establish contact between the pen and the tablet before starting the written production of the word. Whenever the participants did not recognize the object or did not know the name to use to refer to it, they had to put a cross on the line. Each time they thought they were in a tip-of-the tongue state, they had to write down ‘MBL’, which is the French acronym for tip-of-the tongue. At the beginning of each experimental naming session, the participants were given a training phase with a set of 10 items.

Each trial started with the presentation of a ready signal “+” on the screen for 1500 ms and was immediately followed by the presentation of the picture. The picture remained on the screen until the participant started to write the first letter corresponding to the object name. The trials were triggered by the experimenter in order to leave the necessary time for the participants to write the word.

Results

Twelve items were discarded because they yielded less than 50% of correct trials in at least one picture format after all types of errors had been taken into account. The statistical analyses were performed on both written latencies and error rates.

Written latencies

Following Bonin, Méot, Lagarrigue, & Roux (2015), trials with latencies longer than 4000 ms were excluded from the latency analyses (line drawings [$n = 3$], gray levels [$n = 3$], colorized [$n = 8$]). Then, for each type of pictures, latencies exceeding three standard deviations below or above the participants' means were also discarded (line drawings [$n = 88$], gray levels [$n = 71$], colorized [$n = 90$]). Indeed, we closely adhered to the two-step procedure (thresholds + 3 SD) used by Balota et al. (2007), Ferrand et al. (2010) and in a previous study on written spelling in adults (Bonin, Méot, Lagarrigue, & Roux, 2015). When all sets of criteria were applied, we were left with 167 words with 85.2% of the total number of trials in the line drawings condition, 85.7% in the grayscale condition and 85.5% in the colorized condition.[1]

Table 1 shows the descriptive statistics for the variables included in the analyses, namely name agreement (NA), image agreement (IA), visual complexity[2] (VC), imageability,[3] adult word frequency, child word frequency, length (number of letters), orthographic neighborhood (Old20), initial phonology-to-orthography consistency (PO) and bigram frequency. The values of the NA, Imageability and PO consistency variables were mostly at the upper levels of the scales, with relatively marked negative skewness. Name agreement and imageability scores also

[1] Readers who are not familiar with online studies of written naming may be surprised by the number of trials that were eliminated. However, in written naming studies, it is not uncommon for such a number of trials to be discarded. For example, in a recent study of written naming with line drawings, Bonin et al. (2015) found a percentage of valid trials (i.e., 82.3%) similar to that reported here.

[2] As in the majority of picture naming studies, we decided to consider the subjective visual complexity scores. Visual complexity can also be measured in an objective way, i.e., by using file size (Szekely & Bates, 2000). This latter measure is, however, a somewhat raw measure of visual complexity. Some readers may nevertheless ask themselves whether substituting subjective VC for objective VC scores brought about any major changes in the findings. The answer is that they did not.

[3] We chose imageability as a semantic variable because there is a strong consensus among researchers in assuming that this variable reliably indexes semantic code activation (e.g., Yap & Pexman, 2016). Nevertheless, we are aware of other measures of semantics that have been used in English studies (e.g., number of associates, number of semantic neighbors, number of semantic features, see Pexman et al., 2008). However, and unfortunately, such a range of semantic norms is not available in French. We shall return to the influence of semantic richness in picture naming in the "General discussion" section.

Table 1 Statistical characteristics of the stimuli

	Min–max	Mean	SD	Skewness
Name agreement				
Line drawings	75–100	96.2	5.8	−1.87
Grey levels	68–100	94.8	7.4	−1.86
Colorized	75–100	96.8	5.4	−2.12
Image agreement				
Line drawings	1.64–4.82	3.83	.6	−.88
Grey levels	1.75–4.92	3.86	.7	−.83
Colorized	1.83–5.00	3.81	.7	−.38
Visual complexity				
Line drawings	1.00–482	2.72	1.1	.11
Grey levels	1.06–4.88	2.87	1.0	−.05
Colorized	1.00–4.65	2.68	1.0	.17
Imageability	3.56–5.00	4.60	.28	−1.55
Adult word frequency	0–570	36.1	69.6	4.46
Child word frequency	0–927	96.03	141	3.16
Length	3.0–12.0	6.25	1.87	.59
Old20	1.00–4.60	1.96	.68	1.54
PO consistency	.17–100.00	79.48	30.3	−1.32
Bigram frequency	546–21,345	9802	4467	.42

SD standard deviations, *old20* orthographic neighborhood, *PO* phonology-to-orthographic; information about the sources of the norms are provided in the main text

exhibited small variations. The opposite was true concerning the two word frequency variables and the orthographic neighborhood variable.

In order to reduce the skewness of the word frequency measures, all frequencies were log-transformed. In the analyses reported below, word frequency was operationalized as "cumulative frequency".[4] Cumulative frequency is the sum of the z-scores associated with two measures of frequency (in log), namely adult word frequency measures taken from LEXIQUE (New, Pallier, Brysbaert, & Ferrand, 2004) and child frequency taken from MANULEX (Lété et al., 2004). As far as MANULEX child frequency measures are concerned, we used the cumulative frequency corresponding to all grades (G1–5). AoA subjective ratings were not used to investigate age-limited learning effects. We chose to use frequency trajectory instead of rated AoA because frequency trajectory is less correlated with other (subjective) psycholinguistic variables (see Bonin, Barry, Méot, & Chalard, 2004, for a thorough discussion). Frequency trajectory corresponds to the difference between the z-scores associated with the two measures of frequency (LEXIQUE minus MANULEX). Three aspects are worth noting about these scores: (1) We used z-scores and not the raw frequencies (log-transformed) because the LEXIQUE and MANULEX corpora are not the same, and the use of raw frequencies might have introduced discrepancies between the two measures of word frequency; (2) The cumulative frequency and frequency trajectory scores corresponded to the two first

[4] When the analyses were run with adult frequency instead of cumulative frequency, the same pattern of findings as those reported in the "Results" section was found.

factors of the principal component analysis performed on the two frequency measures. As a result, they are uncorrelated, thus permitting more reliable estimations of their effects; (3) The logic of the interpretation of frequency trajectory is comparable with that of AoA ratings: As higher values express low-to-high frequency trajectory, the corresponding words should be learned later when cumulative frequency is controlled for.

The correlations between the different predictors can be seen in Table 2.

There are three main aspects of note. First of all, the correlations between the NA scores of the three different picture formats were all positive (but low). Second, there was a relatively high correlation between IA and PO consistency. Finally, the correlation between number of letters and Old20 was high, suggesting that care should be taken when assessing their respective influence in naming speed in analyses treating them as independent variables.

The mean latency was the longest with the line drawings ($m = 1089$ ms; $sd = 335$; $min–max = 383–3181$) and the shortest with the colorized pictures ($m = 1068$ ms; $sd = 310$; $min–max = 415–2571$), with the grayscale pictures situated between the two ($m = 1076$ ms; $sd = 322$; $min–max = 400–2747$). However, the differences between the three picture formats were somewhat small since the maximum difference was 21 ms.

The correlations between written latencies and the psycholinguistic variables are reported in Table 3. Within each picture format, shorter latencies were associated with higher values for imageability, cumulative frequency, name agreement and, to a lesser extent, image agreement. The reverse was observed for frequency trajectory and visual complexity (but marginally so for the latter variable). Length and Old20 were associated with longer latencies in the grayscale and colorized conditions.

Within-picture format linear mixed model (LMM) analyses

A first series of picture format-specific LMM analyses was run with participants and items treated as random factors which served as the basis for the intercept adjustments in accordance with the mixed model procedure set out in SPSS 22. The predicted scores were the by-format written latencies. The predictors included for each picture format were NA, IA, VC, imageability, cumulative frequency, frequency trajectory, number of letters, orthographic neighborhood (Old20), PO consistency and bigram frequency (in log). Since the ratings corresponding to NA, IA and VC were specific to each picture format, they were standardized within formats. The remaining predictors were standardized for the entire set of stimuli.

In all picture formats, the overall explanation was roughly similar (Table 4). We found reliable effects of NA, imageability and cumulative word frequency in the three picture formats and the effects were all facilitatory, that is to say that the latencies were shorter with increasing values on the corresponding dimensions (e.g., shorter latencies with items having higher word frequency values), the effect of frequency trajectory was also significant and inhibitory in all three formats, with the result that object names having low-to-high frequency trajectory scores took longer to prepare than those having high-to-low scores. Despite having the same signs for all three types of pictures, the effects of image agreement and of visual complexity

Table 2 Correlations among the variables

		Name agreement		Image agreement			Visual complexity			Imageability	Cumulative frequency	Frequency trajectory	Length	Old20	PO consistency	Bigram frequency
		G	C	L	G			G	C							
Name agreement	L	.19	.24	−.01	−.02	.07	.08	.11	.11	.26	.02	−.06	.06	.10	−.00	.06
	G		.25	.03	.11	.06	−.07	−.03	−.04	.23	.12	.02	−.04	.02	−.03	−.01
	C			−.01	.01	.00	.06	.06	.06	.27	.19	−.11	−.12	−.12	.06	.12
Image agreement	L				.79	.78	−.09	−.06	.00	.10	−.21	−.24	.09	.09	−.23	.09
	G					.83	−.12	−.09	−.02	.15	−.21	−.18	.08	.08	−.35	.07
	C						−.13	−.13	−.08	.10	−.30	−.11	.09	.10	−.34	.09
Visual complexity	L							.95	.94	.00	−.02	−.23	−.00	.15	.05	−.06
	G								.95	.07	.03	−.27	−.00	.15	.06	−.05
	C									.07	−.02	−.30	.01	.17	.00	−.05
Imageability											.31	−.22	.05	.04	−.03	.09
Cumulative frequency												.00	−.37	−.38	.32	.07
Frequency trajectory													.00	−.08	−.05	.02
Length														.82	−.15	.20
Old20															−.21	−.02
PO consistency																−.02

With 167 scores, correlations are significant at $p < .001$ if $|r| > .25$, $p < .01$ if $|r| > .19$, $p < .05$ if $|r| > .15$, $p < .10$ if $|r| > .12$; *L* line drawings, *G* grey levels, *C* colorized, *PO* phonology-to-orthography

Table 3 Correlations between written latencies and the predictor variables

	Name agreement	Image agreement	Visual complexity	Imageability	Cumulative frequency	Frequency trajectory	Length	Old20	PO consistency	Bigram frequency
Line drawings	−.1***	−.05**	.05**	−.14***	−.11***	.06***	.00	.02	−.01	−.02
	−.3***	−.14†	.13†	−.41***	−.29***	.18*	.00	.05	.00	−.07
Grey levels	−.12***	−.03	.00	−.12***	−.10***	.05**	.03*	.04**	−.04**	−.00
	−.41***	−.09	.02	−.39***	−.32***	.16*	.13	.16*	−.10	−.03
Colorized	−.11***	−.06***	.04**	−.14***	−.11***	.08***	.03*	.04**	−.02	−.01
	−.35***	−.19*	.11	−.43***	−.33***	.23**	.11	.12	−.05	−.03

*** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$; Values in the first raw are by-trial correlations and values in the second raw are the by-item correlations; *PO* phonology-to-orthography

Table 4 Effects obtained in the within formats LMM with participants and items treated as random factors

	SV $R^2 = .596$	Grey $R^2 = .615$	Color $R^2 = .608$
Intercept	1098.8***	1085.3***	1073.5***
Characteristics of the pictures			
Name agreement	−31.5***	−31.9***	−25***
Image agreement	−12.5	−4.7	−20.5**
Visual complexity	17.7*	7.1	21.2**
Semantic variables			
Imageability	−27.7**	−27***	−29.2***
Lexical variables			
Cumulative frequency	−36.5***	−22.4**	−28.6***
Frequency trajectory	19.9*	16*	19.5**
Sublexical variables			
Length	−24.2	−10.1	−4.9
Old20	18.4	18.2	6.6
PO consistency	7.9	−3.5	.3
Bigram frequency	4.6	3.1	5.1

SV Snodgrass and Vanderwart (1980) pictures; *** $p < .001$; ** $p < .01$; * $p < .05$; *PO* phonology-to-orthography

were not consistent across picture formats. The effect of VC was inhibitory but not reliably so with the grayscale pictures. IA had a reliable influency only in the colorized condition. We did not find reliable effects of sublexical variables. This held true when the number of letters or the orthographic neighborhood measure (Old20) were included separately in order to increase the power of the tests.

Between-picture format linear mixed model (LMM) analyses

In order to compare the effects of the different variables across the three picture formats, we analyzed all the latencies using a unique global LMM (the results are provided in Table 5). The independent variables that had been used in the within-format analyses were also included in the model. NA, IA and VC were now standardized across all formats. The LMM also included two dummy independent variables in order to code the picture formats and their interaction terms with the other independent variables. The interaction terms included in the model were limited to the independent variables that were found to be significant in at least one within-format analysis. The reference category for the formats was alternated in order to compare all pairs. As in the previous analyses, random effects were limited to by-participants and by-items intercepts.

The omnibus test comparing the means of the RT between picture formats was significant, $F(2, 11{,}997.48) = 9.73$, $p < .001$. Pairwise comparisons showed that the mean latencies were significantly longer for line drawings than for the other two formats. Even though the estimation was higher for grayscale than for colorized pictures, the difference was not significant. Within each picture format, Imageability, Cumulative frequency and NA all had reliable facilitatory effects, but the interactions between these variables and the picture format were not significant: Picture format × Imageability,

Table 5 Between-formats linear mixed model

	SV $R^2 = 0.542$	Grey $R^2 = 0.543$	Color $R^2 = 0.523$
b			
Intercept	1097.9***	1083.3***	1076.5***
Picture format			
SV		14.6**	21.4***
Grey			6.8
Characteristics of the pictures			
Name agreement	−10.9*	−7.6*	−10.9*
Image agreement	−12.9*	−1.9	−17.3**
Visual complexity	9.1	2.3	12.4
Picture format × Image agreement			
SV		−11†	4.4
Grey			15.4**
Semantic variables			
Imageability	−37.2***	−30.2***	−31.9***
Lexical variables			
Cumulative frequency	−29**	−26**	−29.6***
Frequency trajectory	16.4*	13.6†	18.7*
Sublexical variables			
Length	−12.1		
Old20	14.5		
PO consistency	2.1		
Bigram frequency	3.6		

The table shows the simple effects of the IVs within each format. Given significant omnibus tests, pairwise comparisons between means and image agreement effects are added

As no interaction with formats were included for sublexical variables, only one effect was evaluated

R^2 were computed within Formats as the r square between raw RTs and their predicted values

SV Snodgrass and Vanderwart (1980) pictures; *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$; *PO* phonology-to-orthography

F(2, 12,279.07) = .79; Picture format × Cumulative word frequency, *F*(2, 12,293.55) = .23; Picture format × NA, *F*(2, 11,847.67) = .22. Frequency trajectory and Picture format also did not interact: Formats * Frequency trajectory, *F*(2, 12,276.23) = .45. However, Frequency trajectory had a reliable inhibitory effect in both the line drawings and colored conditions, but the effect was only marginally significant in the grayscale condition. The only significant interaction with Picture format was observed for IA, *F*(2, 12,351.30) = 4.52, $p < .05$. Image agreement had a reliable facilitatory effect in all conditions except in the grayscale condition. Tests of the simple effects performed on IA confirmed this pattern of results: There was a significant difference of IA between the colored and grayscale conditions and a marginally significant difference of IA between the line drawings and grayscale conditions. Finally, neither Visual complexity nor any of the sublexical variables were significant (note that the Picture formats × Visual complexity interaction was also not reliable, *F*(2, 12,346.38) = 1.90).

Errors

The types of errors obtained with each picture format are reported in Table 6.

Table 6 Types of errors as a function of the three picture formats

	PPE	OE	PLUR	CO	TO	LexS	TOT	DK	Total	%Tot	TECH
SV	146	33	22	52	23	147	14	42	479	10.2	141
Grey	154	42	20	53	31	144	13	53	510	10.8	109
Color	151	46	22	47	30	172	18	43	529	11.1	81
Total	451	121	64	152	84	463	45	138	1518	10.7	331

SV Snodgrass and Vanderwart (1980) pictures, *PPE* phonologically plausible errors, *OE* orthographic errors, *PLUR* incorrect plural or singular, *CO* crossing out, *TO* time-out (>5 s), *LexS* Lexical substitution, *TOT* tip of the tongue, *DK* Do not Know Object or Name of the object, *Total* total number of errors; *%Tot* percentage of errors (technical errors excluded)

At a descriptive level, the number of errors was highest for colorized pictures and lowest for the line drawings. However, both Cochran's Q and Pairwise McNemar tests[5] revealed no significant differences between the picture formats. Among the errors, phonologically plausible errors and lexical substitutions were the most frequent. Stuart-Maxwell (see footnote 5) tests revealed no significant differences in the distribution of errors between the different picture format conditions (all p's $> .1$).

General discussion

The aim of the present study was to investigate how different picture formats influence written naming in relation to several psycholinguistic variables since this is particularly informative about the cognitive architecture that underpins this spelling task. In response to the same concepts presented in three sessions, the participants had to produce on the basis of different picture formats used in the individual writing sessions. There were three types of pictures: line drawings, grayscale drawings and colorized drawings. The main findings can be summarized as follows. First of all, we did not find a reliable general influence of picture format at the level of accuracy. Contrary to our expectations, colored pictures did not reliably yield more accurate naming performances than either line drawings or grayscale drawings. Second, the general speed of writing words from pictures was influenced by the different visual formats, with the result that it took less time to prepare to write words from colorized pictures than from line drawings, with the times for the grayscale pictures being situated between the two. However, the absolute increase in naming speed was somewhat small. Rossion and Pourtois (2004) also found that naming was faster with colorized drawings than with either grayscale drawings or line drawings. The difference between the latter two conditions was not significant. However, and most importantly, we investigated whether and how the effects of a number of predictors of naming speed were

[5] Tests were computed with the exclusion of technical errors. As missing values are not admitted, the exclusion was done for pairs of subjects and items for which there was a technical error in at least one format.

modulated by different picture formats since, as explained in the Introduction, this made it possible to obtain information about the temporal dynamics of written naming. In line with the findings obtained in previous written object naming studies (e.g., Bonin et al., 2004, 2015)[6] that have used the SV line drawings, both within- and between-picture format LMM showed that name agreement, objective word frequency (as indexed by cumulative word frequency), frequency trajectory (note that in the between-picture format LMM analysis, this variable was marginally significant in the grayscale condition), and imageability were reliable determinants of written latencies. The effect of image agreement was, however, less stable. In effect, in the within-picture format LMM analysis, it had a reliable influence only in the colorized condition. This is a rather surprising finding since previous written naming studies have all reported a significant effect of IA with the use of SV drawings (Bonin et al., 2002, 2004, 2015). However, in the between-picture format LMM analysis, the effect of IA on naming speed was significant for both line drawings and their colorized versions, and again, not significant with the grayscale drawings. According to Weekes, Shu, Hao, & Hai (2007), image agreement should have an impact only with SV drawings because these are of poorer quality than the colorized version. However, Bakhtiar, Nilipour, & Weekes (2013) found a reliable effect of image agreement with colorized pictures as we found here. We share Bakhtiar et al.'s claim that, perhaps, image agreement effects are in part cultural in that they depend on socio-cultural similarities between the way common objects are depicted (e.g. dress, car) and the corresponding stored visual representations. It is entirely possible that image agreement is not an appropriate way to evaluate the match between grayscale drawings and stored structural representations of objects. However, the most important finding of the present study, and one which has important theoretical implications, is that, if we except the reliable interaction of IA with picture formats, none of the other interactions with picture format reached significance.

In the Introduction, we argued that an important theoretical issue was to determine whether the physical characteristics of the stimuli only influence prelinguistic stages involved in object naming as assumed by serial-discrete models (Levelt et al., 1999) or limited-cascading models (Bonin et al., 2012b), or whether their influence extends beyond these processes by cascading activation to linguistic processes as predicted by full-cascading models (Humphreys, Riddoch, & Quinlan, 1988). Viewed in combination with previous findings reported by Bonin, Roux, Barry, & Canell (2012b) showing that variables affecting picture format and frequency trajectory did not interact, the present findings strongly suggest that activation does not flow in cascade from the perceptual level to linguistic levels involved in written word production. Thus, previous findings in spoken object naming that have long been taken to support a full-cascading view of word production are inconsistent with the current ones (Humphreys et al., 1988). In effect,

[6] Readers may have noticed that the cited written object naming studies were all conducted in French. Although there is no a priori reason to predict that the findings concerning the influence of the predictors will be different across languages, since we are aware of no other studies that have used the mixed-model approach or the multiple regression approach in a language other than French, this remains an empirical issue.

Humphreys et al. (1988) found that structural similarity (a variable indexing the perceptual stage of object naming) interacted with lexical frequency in spoken object naming, a finding strongly suggesting that there is cascading from object recognition to the phonological level. It is difficult to argue that the conclusions about cognitive architectures are different just because different modalities were investigated in the two studies since there is now reliable evidence suggesting that the two production modes share similar processes and representations (Bonin & Fayol, 2000; Perret & Laganaro, 2012). As pointed out by Levelt et al. (1999), the interaction reported by Humphreys et al. (1988) is problematic because their items confounded word frequency with conceptual familiarity, a variable that is likely to have its locus at the semantic level. Moreover, the interaction between structural similarity and word frequency was not replicated by Snodgrass and Yuditsky (1996).

In this study, we explored the influence of one type of richness on written word production, namely richness of the surface detail, and we failed to find that it exerts an influence beyond prelinguistic levels. (It should be noted that given that visual complexity is also a surface variable, we tested whether this variable interacted with the linguistic variables but we did not find any reliable interactions.) However, we think that there are good reasons to study other types of richness and their potential influence on linguistic processes involved in conceptually-driven production, such as semantic richness. In the language production literature, only a few studies have addressed this issue. Bormann (2011) investigated the number of semantic neighbors in picture naming in adults but failed to find a reliable influence of this factor. More recently, Rabovsky, Schad, and Abdel Rahman (2016) investigated the influence of semantic richness and semantic density in spoken word production. It is important to stress that research on the impact of conceptual/semantic richness is very important since it has considerable theoretical implications for the mechanisms involved in word selection. Word selection is a mandatory process in language production. However, the question of exactly how word selection takes place is currently a matter of debate. Rabovsky et al. (2016) found that pictures depicting objects with a larger number of semantic features had a facilitatory effect on naming times and accuracy. This is presumably because they enhance activation at the conceptual level, this increasing the activation flow to the corresponding lexical representation, and consequently facilitating the naming responses. By contrast, intercorrelational feature density inhibited naming, presumably due to lexical competition.

A future avenue of investigation would be to study the determinants of written naming speed (and accuracy) in response to photographs of objects (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010; Salmon, McMullen, & Filliter, 2010). Few such databases are available at present even though photos of objects are often thought to be more ecologically valid (Brodeur et al., 2014; Moreno-Martínez & Montoro, 2012). However, given the idiosyncrasic nature of individual photographs of objects, it is quite conceivable that written naming performance would not necessarily be better with the use of photographs of objects than it is with the corresponding colorized or black-and-white drawings. (However, Salmon, Matheson, & McMullen's (2014) recent study reported a picture format by manipulability interaction showing that manipulable objects were identified faster when shown as photographs whereas non-manipulable objects were identified equally quickly when depicted as photographs versus line-drawings.) Indeed, some findings in the literature indicate that spoken naming accuracy is lower in response to photographs than in response to the corresponding drawings of the same objects. For instance, the overall name agreement is higher in the SV database (88%), as well as in the Bonin, Peereman, Malardier, Méot, & Chalard (2003) database that was designed to complement the SV set (77%), than in the BOSS database (65%), which contains photographs of objects. As claimed by O'Sullivan, Lepage, Bouras, Montreuil, & Brodeur (2012), photographs include detail and color that prompt participants to rely on idiosyncratic features and, as a result, elicit a number of different names. It is possible that the dynamics of object naming can be investigated better on the basis of prototypical schematic representations of objects, such as line drawings.[7]

Finally, from a methodological point of view, our findings have shown that black-and-white drawings and colorized drawings are equally able to capture the most important determinants of naming speed, with the latter slightly increasing the overall naming speed. Grayscale line drawings should not be recommended for the design of real-time picture naming studies since they provide less stable findings and, in particular, their use does not make it possible to reveal any reliable effects of image agreement on naming speed. These pictures may perhaps be somewhat atypical ways to represent object concepts.

References

Alario, F.-X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments & Computers, 31,* 531–552.

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods Instruments & Computers, 36,* 140–155.

Bakhtiar, M., Nilipour, R., & Weekes, B. S. (2013). Predictors of timed picture naming in Persian. *Behavior Research Methods, 45,* 834–841.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., et al. (2007). The English Lexicon project. *Behavior Research Methods, 39,* 445–459.

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50,* 456–476.

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology, 93,* 89–114.

Bonin, P., & Fayol, M. (2000). Writing words from pictures: What representations are activated and when? *Memory & Cognition, 28,* 677–689.

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways? *Quarterly Journal of Experimental Psychology, 68,* 1268–1294.

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual

[7] The determinants of photos (spoken) naming of objects have been found to be similar to those obtained with the use of colored or black-and-white line drawings (Shao & Stiegert, 2016). However, one very surprising finding in Shao and Stiegert's (2016) study was that rated AoA was not a reliable determinant of naming speed.

familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments & Computers, 35,* 158–167.

Bonin, P., Roux, S., & Barry, C. (2012a). Translating nonverbal pictures into verbal word names: Understanding lexical access and retrieval. In V. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (Chapter 13, pp. 315–332). New York: Psychology Press.

Bonin, P., Roux, S., Barry, C., & Canell, L. (2012b). Evidence for a limited-cascading account of written word naming. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38,* 1741–1758.

Bormann, T. (2011). The role of lexical-semantic neighborhood in object naming: Implications for models of lexical access. *Frontiers in Psychology, 2*, 127. doi:10.3389/fpsyg.2011.00127.

Bramão, I., Faísca, L., Petersson, K. M., & Reis, A. (2010). The influence of surface color information and color knowledge information in object recognition. *The American Journal of Psychology, 123,* 459–468.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE, 5,* e10773.

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of standardized stimuli (BOSS) phase II: 930 new normative photos. *PLoS ONE, 9,* e106953.

Brodeur, M., Kehayia, E., Dion-Lessard, G., Chauret, M., Montreuil, T., Dionne-Dostie, E., et al. (2012). The bank of standardized stimuli (BOSS): Comparison between French and English norms. *Behavior Research Methods, 44,* 961–970.

Brodie, E. E., Wallace, A. M., & Sharrat, B. (1991). Effect of surface characteristics and style of production on naming and verification of pictorial stimuli. *The American Journal of Psychology, 104,* 517–545.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., et al. (2010). The French lexicon project: Lexical decision data for 38, 840 French words and 38, 840 pseudowords. *Behavior Research Methods, 42,* 488–496.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313–338.

Heuer, S. (2016). The influence of image characteristics on image recognition: A comparison of photographs and line drawings. *Aphasiology, 30*, 943–961.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5,* 67–104.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments & Computers, 36,* 156–166.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

Moreno-Martínez, F. J., & Montoro, P. R. (2012). An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS ONE, 7,* e37527.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments & Computers, 36,* 516–524.

O'Sullivan, M., Lepage, M., Bouras, M., Montreuil, T., & Brodeur, M. B. (2012). North-American norms for name disagreement: Pictorial stimuli naming discrepancies. *PLoS ONE, 7,* e47802.

Peereman, R., Lété, B., & Sprenger-Charolles, L. (2007). Manulex-infra: Distributional characteristics of grapheme–phoneme mappings, infra-lexical and lexical units in child-directed written material. *Behavior Research Methods, 39,* 593–603.

Perret, C., & Laganaro, M. (2012). Comparison of electrophysiological correlates of writing and speaking: A topographic ERP analysis. *Brain Topography, 25,* 64–72.

Pexman, P. M., Hargreaves, I. A., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review, 15,* 161–167.

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition, 146,* 240–244.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33,* 217–236.

Salmon, J. P., Matheson, H. E., & McMullen, P. A. (2014). Photographs of manipulable objects are named more quickly than the same objects depicted as line-drawings: Evidence that photographs engage embodiment more than line-drawings. *Frontiers in Psychology, 5,* 1187. doi:10.3389/fpsyg.2014.01187.

Salmon, J. P., McMullen, P. A., & Filliter, J. H. (2010). Norms for two types of manipulability (graspability and functional usage), familiarity, and age of acquisition for 320 photographs of objects. *Behavior Research Methods, 42,* 82–95.

Shao, Z., & Stiegert, J. (2016). Predictors of photo naming: Dutch norms for 327 photos. *Behavior Research Methods, 48*, 577–584.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for names agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments & Computers, 28,* 516–536.

Szekely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter, 12,* 1–33.

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Processes, 10,* 335–342.

Tsaparina, D., Bonin, P., & Méot, A. (2011). Russian norms for name agreement, image agreement for the colorized version of the Snodgrass and Vanderwart pictures and age of acquisition, conceptual familiarity, and imageability scores for the modal object names. *Behavior Research Methods, 43,* 1085–1099.

Weekes, B., Shu, H., Hao, M., & Hai, L. N. (2007). Predictors of timed picture naming in Chinese. *Behavior Research Methods, 39,* 335–342.

Yap, M. J., & Pexman, P. M. (2016). Semantic richness effects in syntactic classification: The role of feedback. *Frontiers in Psychology, 7,* 1394.