

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Sensorimotor and interoceptive dimensions in concrete and abstract concepts

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Villani, C., Lugli, L., Liuzza, M.T., Nicoletti, R., Borghi, A.M. (2021). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. JOURNAL OF MEMORY AND LANGUAGE, 116, 1-12 [10.1016/j.jml.2020.104173].

Availability:

This version is available at: <https://hdl.handle.net/11585/803231> since: 2021-02-22

Published:

DOI: <http://doi.org/10.1016/j.jml.2020.104173>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Caterina Villani, Luisa Lugli, Marco Tullio Liuzza, Roberto Nicoletti, Anna M. Borghi,
Sensorimotor and interoceptive dimensions in concrete and abstract concepts,
«Journal of Memory and Language», Volume 116, 2021, 104173.

The final published version is available online at:

<https://doi.org/10.1016/j.jml.2020.104173>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Running head: SENSORIMOTOR AND INTEROCEPTIVE DIMENSIONS IN CONCRETE AND
ABSTRACT CONCEPTS

Sensorimotor and Interoceptive dimensions in concrete and abstract concepts

Caterina Villani^{a *}, Luisa Lugli^a, Marco Tullio Liuzza^b, Roberto Nicoletti^a, Anna M. Borghi^{c,d *}

^a Department of Philosophy and Communication, University of Bologna, via Azzo Gardino, 23,
Bologna, 40122, Italy.

^b Department of Medical and Surgery Sciences, University of Catanzaro, Viale Europa (Loc.
Germaneto), Catanzaro, 88100, Italy.

^c Department of Dynamic and Clinical Psychology, Sapienza University of Rome, Via degli Apuli
1, Roma, 00185, Italy.

^d Institute of Cognitive Sciences and Technologies, Italian National Research Council.

*Corresponding Authors:

anna.borghi@uniroma1.it (Anna M. Borghi)

caterina.villani6@unibo.it (Caterina Villani)

Abstract

Recent theories propose that abstract concepts, compared to concrete ones, might activate to a larger extent interoceptive, social and linguistic experiences. At the same time, recent research has underlined the importance of investigating how different sub-kinds of abstract concepts are represented. We report a pre-registered experiment, preceded by a pilot study, in which we asked participants to evaluate the difficulty of 3 kinds of concrete concepts (natural objects, tools, and food concepts) and abstract concepts (Philosophical and Spiritual concepts, PS, Physical Space Time and Quantity concepts, PSTQ, and Emotional, Mental State and Social concepts, EMSS). While rating the words, participants were assigned to different conditions designed to interfere with conceptual processing: they were required to squeeze a ball (hand motor system activation), to chew gum (mouth motor system activation), to self-estimate their heartbeats (interoception), and to perform a motor articulatory task (inner speech involvement). In a control condition they simply rated the difficulty of words. A possible interference should result in the increase of the difficulty ratings. Bayesian analyses reveal that, compared to concrete ones, abstract concepts are more grounded in interoceptive experience and concrete concepts less in linguistic experience (mouth motor system involvement), and that the experience on which different kinds of abstract and concrete concepts differs widely. For example, within abstract concepts interoception plays a major role for EMSS and PS concepts, while the ball squeezing condition interferes more for PSTQ concepts, confirming that PSTQ are the most concrete among abstract concepts, and tap into sensorimotor manual experience. Implications of the results for current theories of conceptual representation are discussed.

Keywords: abstract and concrete concepts - mouth motor system - hand motor system – interoception - motor interference – embodied and grounded cognition

Sensorimotor and Interoceptive dimensions in concrete and abstract concepts

Introduction

Categorizing objects and entities in the physical and social environment is fundamental for the survival of our species: categorization helps us to collect information on the world and to simplify its structure forming categories that include similar members, to predict what behavior to expect from different objects/entities, to anticipate how to interact with them etc. Concepts, i.e., the “glue” that link our past, present and future experience (Murphy, 2002), have been broadly distinguished into two main groups, i.e., concrete and abstract ones (e.g. “table” vs. “cause”). Here we do not assume a marked distinction between concrete and abstract concepts (Barsalou et al., 2018); concrete and abstract concepts can be seen more as points in a multidimensional space, the sub-kinds of which can be quite distant from each other (Crutch et al., 2013; Villani, Lugli, Liuzza & Borghi, 2019).

Compared to concrete concepts, abstract concepts have more heterogeneous members and do not possess a single object/entity as referent; they are also more detached from perceptual modalities (Barsalou, 2003), more variable both within and across participants (Borghi & Barsalou, 2014) and more flexible, since they vary more across contexts and situations (Falandays & Spievey, 2019). Previous works revealed higher contextual flexibility for abstract than concrete concepts. For example, Hoffman and colleagues (2013) found substantial variation across words in semantic diversity (*SemD*), which measures the degree of context-dependent variability in word meaning. Concrete concepts appeared in a restricted, inter-related set of contexts and consequently had low semantic diversity values; while abstract concepts tend to be used in a broad range of contexts and consequently showed high values in semantic diversity (see also Hoffman et al., 2016).

According to recent Multiple Representation Views, abstract concepts activate the sensorimotor system but also the emotional dimension (e.g. Newcombe et al., 2012; Kousta et al., 2011; Vigliocco et al., 2014), and the linguistic and social one (Borghi et al., 2019a; Dove, 2019; Glenberg, 2019). Here we focus on the WAT (Words As social Tools) theory (Borghi & Cimatti, 2009; Borghi &

76 Binkofski, 2014; Borghi et al., 2018a, 2019a, 2019b), according to which abstract concepts are not
 77 only grounded in sensorimotor experience, similarly to concrete concepts, but activate linguistic,
 78 social and interoceptive experiences to a larger extent than concrete concepts. The WAT view
 79 proposes that the activation of linguistic and social experience during processing and use of abstract
 80 concepts might be due to different reasons. These reasons are not incompatible, and they all stem
 81 from the basic consideration that abstract concepts are more complex than concrete ones. We consider
 82 abstract concepts as more complex than concrete ones for a variety of reasons. First, they generally
 83 do not have a single object as referent, as concrete concepts, but refer to more complex scenes and
 84 elements. For example, the abstract concept of “cause” involves an agent, a patient, an action
 85 (Pulvermüller, 2018). Second, they are more complex to learn because their members are more
 86 heterogeneous and different than those of concrete concepts - justice situations are certainly more
 87 diverse from each other than different tables. This complexity has a behavioral effect, i.e. the widely
 88 replicated concreteness effect (e.g. Paivio, 1990). Abstract concepts require generally longer times to
 89 be processed, and are recalled less accurately than concrete concept.

90 The first reason for the importance of the social and linguistic dimension for abstract concepts is
 91 their particular acquisition modality: linguistic inputs offered by others are crucial in order to keep
 92 together the variety of heterogeneous events and situations that characterize abstract concepts (labels
 93 as glue of heterogeneous experiences) (Lupyan, 2019). During abstract concepts processing
 94 participants might re-enact such verbal linguistic acquisition experience. Even if further research
 95 should clarify this, this mechanism might be present also when words are in the written modality,
 96 influential especially for learning low-frequency abstract words. Indeed, evidence suggests (e.g.
 97 Topolinski & Strack, 2009) that during reading we simulate the motor responses associated with
 98 verbal stimuli. The second reason and the third reason stem from the feeling of uncertainty and the
 99 metacognitive awareness that our knowledge of abstract concepts is scarce and inadequate (see
 100 Borghi, Fini & Tummolini, under review). This awareness might lead to two different outcomes. The
 101 first is the need to rehearse and re-explain to ourselves the word meaning, possibly through inner

102 speech. The second is the preparation to ask information to competent others (social metacognition;
 103 Borghi et al., 2018a; see also Shea, 2018; Prinz, 2014). Importantly, all these mechanisms might not
 104 only lead to the activation of linguistic and social networks, but also engage the mouth motor system
 105 more than processing of concrete concepts does. In line with an embodied account, we namely
 106 hypothesize that using both overt and inner speech implies a motor simulation that involves the mouth
 107 motor system (Topolinski & Strack, 2009; Alderson-Day, B., & Fernyhough, 2015). Consistently, a
 108 variety of studies have demonstrated that the mouth motor system is involved to a larger extent during
 109 abstract than during concrete concepts processing (review in Borghi et al., 2019a), and in particular
 110 during processing of mental states abstract concepts (Dreyer & Pulvermüller, 2018; Ghio et al., 2013).
 111 Furthermore, it is possible that "concrete" concepts may be more readily referenced through non-
 112 verbal/non-linguistic means e.g., deictic gestures, as they more likely refer to physical objects in
 113 space, while "abstract" concepts may need to be supplemented by other communicative tools (such
 114 as inner speech).

115 An important development in recent literature on abstract concepts relates to the recognition that
 116 they are not a unitary whole, but that subtypes of abstract concepts exist (Desai et al., 2018; Fischer
 117 & Shaki, 2018; Fingerhut & Prinz, 2018; Villani et al., 2019). In the domain of concrete concepts,
 118 instead, much research on sub-kinds of concepts has been conducted. Neuropsychological and brain
 119 imaging studies have focused in particular on the double dissociation between living and non-living
 120 entities and on their different neural representation (Warrington & Shallice, 1984; review: Forde &
 121 Humphreys, 2005), behavioral studies have investigated the roughly correspondent distinction
 122 between artifacts and natural objects and on how it develops in children (Keil, 1989). In the last few
 123 years there is growing interest for concepts such as food, that is for concepts that are neither artifact
 124 nor natural but that can be both depending on the circumstances (Rumiati & Foroni, 2016).

125 Our study aims to investigate the fine-grained differences in the representation of abstract and
 126 concrete concepts and to identify possible sub-kinds of both kinds of concepts. Building on previous

127 studies (see below), we decided to use the same rating task: in the Pilot study, we asked participants
128 to rate the difficulty and the pleasantness of different abstract words; in the Experiment, we asked
129 participants to rate the difficulty of both concrete and abstract words. Crucially, participants were
130 assigned to different conditions that were supposed to interfere with a specific kind of concept, thus
131 to increase the perceived difficulty of specific kinds of words.

132 Pilot study

133 The current study builds on the method of a previous study in preparation (Borghi & Lugli, in prep;
134 Lugli & Borghi 2017) and for the selection of materials on a recently published norming study (Villani
135 et al., 2019).

136 In the study by Borghi and Lugli participants of different groups were asked to rate the degree of
137 pleasantness and difficulty of concrete and abstract concepts while performing a concurrent task.
138 Participants were told that their evaluations would be used to contribute to select the verbal stimuli
139 for an experiment, and were asked to what extent they perceived the presented words as difficult and
140 pleasant, without any further specification. We chose to avoid orienting participants toward a specific
141 meaning of difficulty, and to use the common sense of the word. However, we think that the cover
142 story leads them to interpret difficulty in terms of “difficulty in processing”. Participants were
143 assigned to 3 different conditions: in the ball condition they had to rhythmically squeeze a ball, in the
144 gum condition to rhythmically chew gum, and in the candy condition to suck a candy. These
145 conditions were designed to verify whether actively moving the mouth interfered with abstract
146 concepts processing, and actively manipulating a ball with processing of concrete concepts. The
147 candy condition was intended as a control one. A higher processing difficulty should lead to an
148 increase in rated difficulty and a decrease in rated pleasantness.

149 The rationale of our pilot experiment builds on this previous work, but with two important differences.
150 First, we intended to test not only the effect of the mouth active movement (gum chewing) and of the
151 hand active movement (ball squeezing) on difficulty and pleasantness ratings, but also the effects of

152 interoceptive experience (Connell et al., 2018; Borghi et al., 2019a) and of social experience (Borghi
 153 & Cimatti, 2009; Borghi & Binkofski, 2014; Borghi et al., 2018a, 2019a) on abstract concepts
 154 processing. Hence, in the Pilot study we added to the gum and to the ball condition two further
 155 conditions, i.e., the interoceptive condition, in which participants were asked to hold an instant cold
 156 or warm pack, and the social condition, in which they were required to hold the hand of a confederate.
 157 Second, the main aim of the Pilot study was not to identify differences between abstract and concrete
 158 concepts, but more subtle differences within abstract concepts. To identify sub-kinds of abstract
 159 concepts, we relied on the study by Villani et al. (2019). In this norming study participants were asked
 160 to evaluate 425 Italian abstract words on 15 dimensions (i.e., Abstractness, Concreteness,
 161 Imageability, Context availability, Body-Object-Interaction, Modality of Acquisition, Age of
 162 Acquisition, Perceptual modality strength, Metacognition, Social metacognition, Interoception,
 163 Emotionality, Social valence, Hand and Mouth activation). We then performed a cluster analysis that
 164 led to the identification of 4 clusters of abstract concepts, i.e., Philosophical and Spiritual concepts
 165 (PS) (e.g., value, belief), Emotional and Mental State concepts (EMS) (e.g., anger), Social and Self
 166 concepts (SS) (e.g., kindness) and Physical Space Time and Quantity (PSTQ) (e.g., reflex, sum). PS
 167 concepts were more abstract than the others, i.e., acquired late (e.g. Kuperman et al., 2012) and
 168 through language, and more characterized by the tendency to ask the meaning to others (social
 169 metacognition), PSTQ concepts were more concrete, i.e., more imageable, more characterized by
 170 bodily interactions with the environment. SS and EMS were more characterized by inner grounding,
 171 i.e. interoception and emotional valence and metacognition, and by sensorimotor properties (taste,
 172 smell, etc.). Further details of four kinds of abstract concepts and their cluster distributions can be
 173 found at <https://osf.io/4bztv/>. As in the previous study by Borghi and Lugli (in prep.), participants
 174 were required to perform pleasantness and difficulty judgments on a 5-point scale. Both scores and
 175 response times were recorded. The reason why we choose to use pleasantness and difficulty ratings
 176 is due to the fact that, in the literature, a relationship has been found between abstraction and
 177 disfluency, and concreteness and fluency (Alter & Hoppeneier, 2008, but see one experiment for a

178 failure to replicate). Increased fluency augments preference for a given stimulus (Winkielman et al.,
 179 2003). For example, the increased fluency of pronunciation simulation, owing to the exposure, leads
 180 to an increase of word pleasantness (Topolinski & Strack, 2009).

181 Participants were assigned to 4 different conditions: ball condition (they were asked to rhythmically
 182 squeeze a softball), interoceptive condition (they were asked to hold an instant cold or warm pack);
 183 social condition (they were asked to hold the hand of a confederate); gum condition (they had to
 184 rhythmically chew gum). We predicted that judgments of difficulty would increase in the ball
 185 condition more with the more concrete PSTQ concepts than with the other abstract concepts, that the
 186 interoceptive condition would lead to an increase of difficulty and a decrease of pleasantness ratings
 187 especially with EMS and SS concepts, which are more directly related to social and emotional aspects,
 188 that the social condition would lead to an interference mostly with SS concepts, and that the gum
 189 condition would interfere mostly with judgements produced in the most abstract PS concepts.

190 Method

191 Participants

192 129 students (102 female, 18 left-handed; $M_{age} = 24.2$, $SD_{age} = 3.7$) of the University of Bologna
 193 participated voluntarily. All participants were recruited among the students of a Psycholinguistic
 194 course. They were randomly assigned to the four conditions, resulting in 30 participants for ball
 195 condition, 39 for interoceptive condition, 26 for social condition and 34 for gum condition. All
 196 participants assigned to each condition were tested together in a room equipped with computers.

197 Materials

198 60 concepts taken from the previously identified four clusters were selected. We considered the most
 199 representative words for each cluster (i.e., the ones with the smallest distance from the centroid; mean
 200 distance = 2.44, max. 6.75; min. 0.72) and selected them for their value of Abstractness in a range
 201 from 1 (less abstract) to 7 (more abstract). Of 60 concepts, 13 were selected from PSTQ cluster

(Mean = 2.72, SD = 0.58), 21 from PS cluster (Mean = 4.96, SD = 0.97), 11 from SS cluster (Mean = 4, SD = 0.78) and 15 from EMS cluster (Mean = 4.29, SD = 0.65).

204

205 Procedure

206 Participants were asked to evaluate on a 5-point Likert scale the difficulty ranging from 1 = “very
207 easy” to 5 = “very difficult” and the pleasantness ranging from 1 = “very unpleasant” to 5 = “very
208 pleasant” of each word presented.

209 Each participant was instructed to provide both difficulty and pleasantness ratings in different blocks;
210 the order of the blocks was counterbalanced across participants. During the evaluation, they had to
211 perform a concurrent task. They were randomly assigned to four different conditions: gum chewing
212 (they were asked to chew gum following the rhythm of a metronome) (Topolinski & Strack, 2009;
213 Topolinski et al., 2014), interoceptive (they were asked to hold an instant cold or warm pack, that
214 kept the temperature until the end of the task), social condition (they were asked to hold the hand of
215 a confederate), ball squeezing (they were required to manipulate a softball following the rhythm of a
216 metronome). The order to the trials was fully randomized, with the exception to not repeat the same
217 word twice in succession.

218

219 Data analysis and results

220 Because of the ordinal nature of the dependent variable (responses on a Likert-type format), we
221 conducted our analyses using Cumulative link mixed models (logit link function) using the clmm
222 function from the ordinal (Christensen, 2018) R library. We modeled participants and words as
223 random intercepts in order to account for the dependence among observations. Ideally, we should
224 have modeled random slopes for each participant and word in order to better control for the Type I
225 error (Barr et al., 2013), but it led to severe convergence issues. RTs were added as a predictor in the

model in order to control for the effect of speed on the pleasantness and difficulty judgments. A Model comparison through Likelihood Ratio Tests was conducted in order to test the overall effects of the Condition, the Cluster, and their interaction.

We did not find any statistically significant effect for either the Condition, the Cluster or their interaction on pleasantness ratings (see Table 1). When analyzing difficulty ratings, we did find a main effect of the cluster (see Table 2). In fact, PS words were more likely to be rated as less difficult as compared to words belonging to other clusters. We did not find any other statistically significant effect for either the Condition or for the Condition x Cluster interaction.

We expected to observe that the interference in the gum chewing condition should be stronger for PS abstract concepts, because of their high level of abstractness. However, the planned contrast on interaction between cluster PS and Condition (gum vs. social, interoceptive and ball in PS clusters > gum vs. social, interoceptive and ball in other clusters) was not significant ($p = .93$).

Table 1. Model comparison of the effects on pleasantness ratings. The table reports a Likelihood ratio test between models where a predictor at time was entered. AIC = Akaike Information Criterion. No.par = number of parameters of the model.

Predictors	No.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
RT	7	16981	-8483	4.37	1	0.037
Condition	10	16984	-8482	2.41	3	0.492
Cluster	13	16989	-8482	1.34	3	0.721
Condition x Cluster	22	17000	-8478	7.03	9	0.634

244 **Table 2.** Model comparison of the effects on difficulty ratings. The table reports a Likelihood ratio
 245 test between models where a predictor at time was entered. AIC = Akaike Information Criterion.
 246 No.par = number of parameters of the model.

Predictors	No.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
RT	7	16980	-8483	20.70	1	< .001
Condition	10	16982	-8481	3.67	3	.300
Cluster	13	16964	-8469	23.61	3	< .001
Condition x Cluster	22	16968	-8462	14.08	9	.120

247

248

249 Experiment

250 Potential problems of the Pilot study were that we had limited ourselves to consider sub-kinds of
 251 abstract concepts, and concrete words were not introduced. In addition, the social manipulation might
 252 have not been successful because touching someone you do not know could render it very difficult to
 253 concentrate on the experiment. Finally, in three of four manipulated conditions participants were
 254 asked to use their hand – this might have reduced the differences between the conditions.

255 The present pre-registered Experiment was designed to overcome these limitations. We confined
 256 ourselves to difficulty rating, for which the results of the previous study were more clear-cut. We
 257 selected three kinds of concrete and abstract concepts, controlled the materials, and modified two of
 258 the four conditions. The conditions to which participants were randomly assigned were: ball
 259 squeezing, gum chewing, heart beating, and articulatory suppression. For the heart beating condition
 260 we asked participants to estimate their heart beat pace and at the end of the task to report if they had
 261 noticed any change; self-estimation of heart beating within a given time is a task often used to measure
 262 interoceptive awareness (Schandry, 1981; Garfinkel et al., 2015). In order to test whether processing

263 of abstract concepts does not only involve the mouth but implies use of inner speech, we introduced
 264 an articulatory suppression condition, since AS is often used to test involvement of inner speech
 265 (Alderson-Day & Fernyhough, 2015). In the articulatory suppression condition participants were
 266 required to rhythmically pronounce the syllable “ba ba ba”. Finally, we introduced a control
 267 condition, in which participants were asked to evaluate the difficulty of the words without performing
 268 any additional task. The control condition was introduced primarily because the conditions might
 269 differ in terms of executive demands. Conditions that capture more attention could more easily lead
 270 to interference, while conditions that involve low processing load might not affect the results (Connell
 271 & Lynott, 2012). We introduced the control condition also to better understand whether an
 272 interference or a facilitation occurred with respect to the baseline. It is worth noting that the control
 273 condition was not present in the original design and in the preregistration; we introduced it because
 274 the reviewers asked for it. Differently from the other conditions, in the control condition participants
 275 were tested online, since the lock-down due to the spread of COVID-19 did not allow us to test
 276 participants in the lab.

277

278 **Hypotheses**

279 Hypothesis 1 (directional). Ball squeezing condition: if processing of more concrete concepts, and
 280 particularly of tools, involves to a larger extent the manual motor system, i) we predicted that this
 281 condition would interfere more with concrete than with abstract concepts, thus increasing the rated
 282 difficulty of the concrete concepts, compared to the other conditions. ii) The interference effect should
 283 be particularly strong for tools, increasing their perceived difficulty, and then for food items. iii)
 284 Within abstract concepts, we intended to explore whether the ball squeezing condition would create
 285 more interference with the more concrete among the abstract concepts, i.e., PSTQ.

286 Hypothesis 2 (directional). Gum chewing condition: if processing of abstract concepts activates the
 287 mouth motor system to a larger extent than processing of concrete concepts, then i) we predicted that

the gum chewing condition would interfere more with abstract concepts than with concrete concepts of animals and tools, leading to an increase in difficulty of more abstract compared to more concrete concepts. Within concrete concepts ii) we predicted that gum chewing would modulate the food items to a larger extent, either determining a decrease or increase of difficulty (facilitation or interference), because of the relationship between food items and mouth motor system.

Hypothesis 3 (directional). Articulatory suppression condition: if processing of abstract concepts not only activates the mouth motor system but specifically involves inner speech, then i) we predicted that the articulatory suppression condition would interfere more with abstract concepts than with concrete concepts, increasing the perceived difficulty of the first with respect to the second, and in particular ii) for the more abstract concepts, i.e., PS.

Hypothesis 4 (directional). Heart beating condition: if processing of abstract concepts not only activates the mouth motor system but also the interoceptive dimension to a larger extent than processing of concrete concepts, then i) we predicted that the heart beating condition would interfere more with abstract concepts than with concrete concepts, increasing the difficulty of the first and reducing that of the second. This should occur in particular with abstract concepts that involve more the emotional and social dimension, i.e., with EMSS (see results by Connell et al., 2018, showing that interoception characterized primarily emotional concepts). Within concrete concepts, ii) we intended to explore whether the heart beating condition would create more interference with the concepts of animals, because of their animacy.

Method

Material selection.

The words were selected from both the database by Della Rosa et al. (2010) and our database (Villani et al., 2019). More specifically, the selection of concrete words was completely based on the database of Della Rosa et al. (2010). Concrete words included 10 natural objects (animals, e.g., lion, camel), 10 manipulable artifacts (tools, e.g., hammer, broom) and 10 food items (e.g., carrot, eggplant)

Concrete stimuli are shown in Table 3. We selected these three categories because these can be considered almost exhaustive of the categorical space and are used in the majority of studies on concrete concepts. Since the seminal work by Warrington and Shallice (1984), many studies on concrete concepts have focused on the distinction between artifacts and natural objects (for a review on the living/nonliving double dissociation see Forde & Humphreys, 2002). Recent studies are targeted at investigating the specificity of food concepts, which possess a special status since they are neither natural nor artifact objects (Rumiati & Foroni, 2016). Within artifacts, we focused on tools, more likely to activate the hand motor system (see Martin, 2007, for a review).

Abstract words were selected taking into consideration the two databases. Abstract words included words present in Della Rosa et al. (2010) but were selected by means of the clusters that emerged in the study by Villani et al. (2019): 10 words were selected from the cluster Philosophical and Spiritual concepts (PS, e.g., destiny, morality), 10 from the cluster Physical Space Time and Quantity (PSTQ, e.g., number, acceleration). Because the differentiation between Emotional and Mental State concepts (EMS, e.g., shame) and Social and Self concepts (SS, e.g., calm) was not clear cut, we decided to collapse the two clusters and selected 10 words from them (5 for each cluster). Abstract stimuli are shown in Table 4. Importantly, the different sub-groups of concrete and abstract words did not differ across main psycholinguistic dimensions, including the number of syllables, familiarity, absolute and relative frequency. Further characteristics of the selected concrete and abstract words in terms of dimensions and psycholinguistic variables are available in an online repository as Supplementary Materials (<https://osf.io/ypx7s/>).

Table 3. Selected concrete words from Della Rosa et al. (2010) database. Frequency values for each word were determined by CoLFIS, a lexical database of written Italian (Bertinetto et al., 2005).

Italian word	English word	Frequency value	Numbers of letters	Frequency absolute mean	N Letters mean
Banana	Banana	24	6		
Carota	Carrot	41	6		
Uva	Grapes	26	3		
Fragola	Strawberry	30	7		
Fungo	Mushroom	38	5		
Melanzana	Eggplant	13	9		
Peperone	Pepper	27	8		
Pomodoro	Tomato	88	8		
Torta	Cake	67	5		
Zucca	Pumpkin	33	5	<i>Concrete Food</i> = 38.7	6.2
Lampada	Lamp	76	7		
Martello	Hammer	26	8		
Scopa	Broom	12	5		
Bottiglia	Bottle	122	9		
Coltello	Knife	117	8		
Trapano	Drill	9	7		
Ombrello	Umbrella	31	8		
Forchetta	Fork	25	9		
Matita	Pencil	45	6		
Pennello	Brush	29	8	<i>Concrete Tool</i> = 49.2	7.5
Cane	Dog	328	4		
Leone	Lion	78	5		
Maiale	Pig	40	6		
Cammello	Camel	15	8		
Pecora	Sheep	56	6		
Mucca	Cow	12	5		
Piccione	Pigeon	19	8		
Gallina	Chicken	32	7		
Pappagallo	Parrot	12	10		
Insetto	Insect	76	7	<i>Concrete Animal</i> = 66.8	6.6

Table 4. Selected abstract words from Della Rosa et al. (2010) and Villani et al. (2019) database. Frequency values for each word were determined by CoLFIS, a lexical database of written Italian (Bertinetto et al., 2005).

Italian word	English word	Frequency value	Numbers of letters	Frequency absolute mean	N Letters mean
Accelerazione	Acceleration	29	13		
Inizio	Beginning	453	6		
Schema	Scheme	116	6		

Area	Area	483	4	
Numero	Number	1196	6	
Risultato	Results	902	9	
Punizione	Punishment	76	9	
Rimedio	Remedy	71	7	
Sforzo	Attempt	258	6	
Denaro	Money	337	6	<i>Abstract PSTQ</i> = 392.1 7.2
Morale	Moral	85	6	
Descrizione	Description	66	11	
Motivo	Motive	602	6	
Salvezza	Salvation	85	8	
Destino	Fate	266	7	
Paradiso	Paradise	92	8	
Enigma	Enigma	20	6	
Peccato	Pity	178	7	
Giudizio	Judgement	371	8	
Logica	Logic	117	6	<i>Abstract PS</i> = 188.2 7.3
Calma	Calm	110	5	
Gioia	Joy	235	5	
Amicizia	Friendship	212	8	
Conflitto	Conflict	186	9	
Gentilezza	Kindness	25	10	
Vendetta	Revenge	112	8	
Ansia	Anxiety	137	5	
Vergogna	Shame	101	8	
Simpatia	Liking	132	8	
Paura	Fear	698	5	<i>Abstract EMSS</i> = 194.8 7.1

347

348

349 **Sample size rationale**

350 We conducted a power analysis through the pwr package in R (Champley, 2018). In order to achieve
351 a power of 80% with a critical alpha of .05 divided by the number of unpaired t-tests (.05/9 = .0055)
352 that would allow us to test our pre-registered hypotheses, and assuming a medium effect size (Cohen's
353 D = 0.5) (Cohen, 1988), and having a directional hypothesis we would need 93 participants per group
354 (total N = 372). Since it would have been unfeasible to achieve that number due to objective
355 constraints (N = 120 students enrolled in the class, and a time limited to one month), we decided to
356 determine an effect size as the minimum amount of observations needed to have a relatively stable
357 estimate. Based on Green (1991)'s rule of thumb for determining the smallest sample size, we would

358 need $104 + k$ (where k is the number of predictors, i.e., number of groups $-1 = 3$). Therefore, any
 359 sample size greater than 107 would be enough to avoid overfitting.

360 However, since inferences based on the Null Hypothesis Significance Testing are problematic without
 361 adequately controlling for the Type I and Type II error at the same time (Dienes, 2008), we used a
 362 Bayesian approach, instead. The sample size consisted of around 100-120 participants (25-30 per
 363 condition).

364

365 **Participants**

366 130 students participated (108 female, 14 left-handed; $M_{\text{age}} = 24$ $SD_{\text{age}} = 2.5$). Participants were
 367 volunteers recruited among the students of a Psycholinguistic course; they were students of the first
 368 or second year of the Master's degree in Semiotics, Philosophy, Italian Studies, Language and
 369 Communication. Each participant was randomly assigned to one of the five groups (gum chewing,
 370 articulatory suppression, heart beating, ball squeezing, control), resulting in 26 participants for each
 371 group. All participants were tested together in a room equipped with computers, except for
 372 participants in control condition who were tested online.

373 **Procedure**

374 Participants were asked to evaluate the difficulty of the stimuli using a 5-point Likert scale where 1
 375 corresponded to "very easy" and 5 to "very difficult". During the evaluation they have to perform a
 376 concurrent task depending on the condition to which they were assigned: they were asked to chew
 377 gum following the rhythm of a metronome (gum chewing), to rhythmically pronounce the syllable
 378 "ba ba ba" (articulatory suppression), to estimate their heart beat pace and in the end of the task report
 379 if they have noticed any change (heart beating), to manipulate a softball following the rhythm of a
 380 metronome (ball squeezing). In the control condition no concurrent task was introduced. In all

381 conditions, the full list of stimuli was presented twice resulting in a total of 120 words. The order to
382 the trials was fully randomized, with the exception to not repeat the same word twice in succession.

383

384 **Data analysis**

385 A detailed pre-registered analytic plan can be found on the Open Science Framework repository at
386 the following link: <https://osf.io/3qu7t> Notice that some of the data were collected prior to pre-
387 registration, even if we have not performed any kind of analysis on them.

388 We measured the evaluations provided on a 5-point scale; we also measured the response times
389 required to respond and consider them as a covariate. Predictors: Modality of Acquisition (MoA,
390 Wauters, 2003), abstractness and concreteness.

391 Given the clustered nature of our design (word categories were manipulated within participants) and
392 to minimize any loss of information, we decided to analyze our data through a multilevel model (also
393 known as mixed models, Pinheiro & Bates, 2000). In this way, we took into account participants and
394 words as sources of variation. To this purpose, we modeled participants' and words' intercepts as
395 random effects (i.e. (1|participant) and (1|word) in Wilkinson notation). Although it is recommended
396 to keep the random structure maximal (Barr *et al.*, 2013), adding the random slopes led to
397 convergence issues, thus we decided to model only the random intercepts.

398 Furthermore, Liddle and Kruschke (2018) have recently demonstrated that treating a response
399 measured at an ordinal level of measurement (e.g., Likert response format) like a variable measured
400 at an interval level can lead to false alarms, misses, and even inversions. For this reason, we followed
401 the recommendations from Buerkner and Vuorre (2019), and modeled our responses within an ordinal
402 model, using a cumulative model with a probit or a logit link function. To decide which link function
403 had better predictive accuracy, we fitted them both and selected the best fitting model in terms of the
404 Watanabe-Akaike information criterion (WAIC; Watanabe, 2010).

405 In the first model we tested whether the difficulty ratings were affected by the interaction between
406 the sub-kinds of concepts and the experimental conditions. We set participant-level and word-level
407 random intercepts in order to account for non-independence among our observations.

408 Furthermore, we conducted our analyses within a Bayesian framework, as it provides more flexibility
409 for parameter estimation, and allows us to make claims on the relative evidence in favor of a
410 hypothesis (e.g., H_1) compared to another (e.g., H_0 , Wagenmakers, 2007).

411 The analysis was conducted in the Bayesian framework provided by the brms (Bayesian regression
412 models using ‘Stan’) library (Bürkner, 2017, 2018) in R. All the models were fit using three different
413 priors on the coefficients, to assess the sensitivity of the analysis: uninformative (flat prior, default in
414 brms), weakly informative (normal distribution centered on zero and with a standard deviation of 5),
415 or a narrower prior (normal distribution centered on zero and with a standard deviation of 1).

416 Our hypotheses were tested through the “hypothesis” function on brms, which assesses the relative
417 strength of evidence in favor of competitive hypotheses using the Savage-Dickey density ratio
418 method, which compares the plausibility of a hypothesis (e.g., the null hypothesis “abstracts =
419 concrete” under the prior vs. under the posterior probability distribution). Bayes factors were reported
420 following the convention of reporting the hypothesis tested as a subscript: BF_{10} stands for relative
421 evidence for the alternative (H_1) vs. the null (H_0), whereas BF_{01} stands for relative evidence for the
422 alternative (H_0) vs. null (H_1). We also sampled from the posterior distribution for computing the
423 posterior probability (PP) of the alternative, directional, hypothesis. We chose the best fitting link
424 function using the WAIC (the least the best).

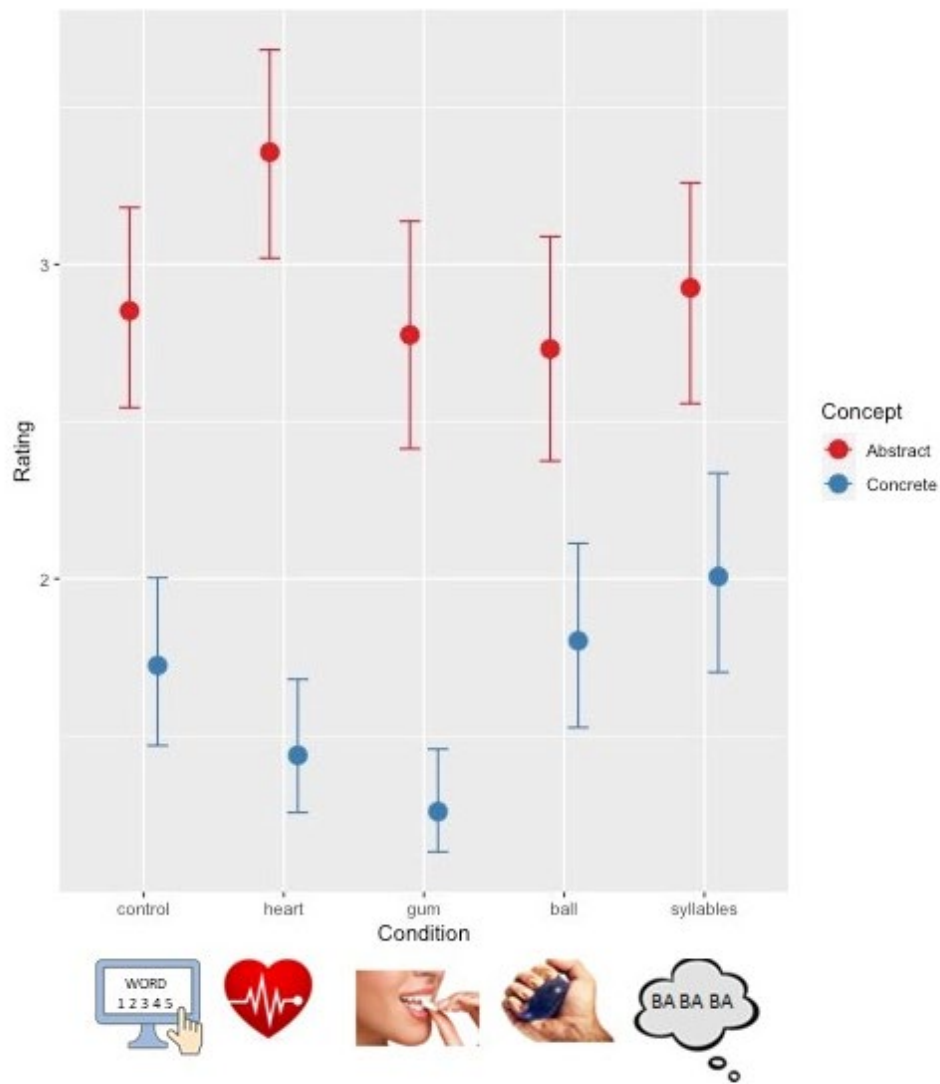
425 We interpreted the relative strength of evidence using the labels provided by Jeffreys (1961, revised
426 by Lee and Wagenmakers, 2013). Furthermore, checking the inclusion of zero within the 95%
427 posterior credible intervals were used as additional information about the plausibility of the null
428 hypothesis (and/or estimates of practical irrelevance) given the data.

429 Since Bayesian Multilevel models are relatively robust to outliers (Nezlek, 2011), especially with a
430 relatively narrow priors as the ones used in our analysis, we did not exclude outliers. We excluded
431 data that was incorrectly entered (e.g., age > 99, Likert scale response > 5, etc.). Missing data were
432 dealt with using a pairwise deletion.

433 **Results**

434 We fit two models containing only the intercepts (fixed and random), changing only the link function
435 for the ordinal cumulative model (logit vs. probit). We found that the ordinal cumulative model with
436 the logit (WAIC = 29266.7) link function outperformed the ordinal cumulative model with the probit
437 link function (WAIC = 29266.7, Δ WAIC = 7.9). We therefore used an ordinal cumulative model with
438 the logit link function for all the following analyses (Table 5).

439 In the first model we modeled the variables just in terms of abstract vs. concrete words and of
440 experimental conditions (Figure 1). The estimates for the model with uninformative and flat priors
441 appeared to lead to similar results, but the narrow priors lead to somewhat more conservative
442 estimates – unsurprisingly. Therefore, we reported the results when placing a narrow prior on the
443 parameters.



444

445 *Fig.1* Interaction plot of ratings mean versus conditions (control, heart, gum, ball, syllables) for
 446 abstract and concrete concepts. Error bars indicate the 95% credible intervals.

447

448

449 **Table 5.** Estimates and 95% posterior credibility intervals (PCIs) for the estimates for the model in
 450 which we tested for the effect of concreteness (abstract vs. concrete) and experimental condition
 451 (control, heart, gum, ball, syllables) using a narrow prior (normal distribution with mean = 0 and SD
 452 = 1). Abstract concepts and heart beating conditions are set as reference variables for the concreteness
 453 and the experimental conditions, respectively. Boldfaced: the estimates whose 95% PCIs do not
 454 include the effect of zero.

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept[1]	-2.72	0.33	-3.38	-2.09
Intercept[2]	-0.62	0.33	-1.27	0.01
Intercept[3]	1.2	0.33	0.54	1.82
Intercept[4]	3.3	0.33	2.65	3.93
Condition Heart	1.01	0.38	0.29	1.76
Condition Gum	-0.16	0.39	-0.91	0.6
Condition Ball	-0.25	0.4	-1.03	0.52
Condition Syllables	0.14	0.38	-0.62	0.89
Concept Concrete	-2.48	0.31	-3.07	-1.85
Condition Heart: Concept Concrete	-1.83	0.11	-2.04	-1.61
Condition Gum: Concept Concrete	-1.35	0.12	-1.58	-1.12
Condition Ball: Concept Concrete	0.44	0.11	0.23	0.65
Condition Syllables: Concept Concrete	0.55	0.1	0.35	0.75

455

456

457 Hypothesis 1. i) We predicted that the ball squeezing condition would have increased the perception
458 of the difficulty of concrete concepts (vs. abstract ones). To test this hypothesis, we tested whether
459 the difference between abstract and concrete concepts in the ball condition was different as compared
460 to other conditions. We found extreme evidence that this difference was smaller in the ball condition,
461 as compared to the control, the gum and the heart beating conditions ($BF_{10s} > 100$, posterior
462 probability (PP) = 100%). However, there was moderate evidence that there was no difference
463 between the difference between abstract and concrete concepts in the ball condition as compared to
464 the articulatory suppression condition ($BF_{01} = 8.88$, PP = 16%). We also tested whether the difficulty
465 ratings for concrete concepts in the ball condition were higher than in other conditions. We found
466 extreme evidence in favor of the hypothesis that difficulty ratings for concrete concepts in the ball
467 condition were higher than in the gum conditions ($BF_{10} > 100$, PPs = 100%), and moderate evidence

that difficulty ratings for concrete concepts in the ball condition were higher than in the heart beating condition ($BF_{10} = 3.4$, $PP = 99\%$). However, there was anecdotal evidence that difficulty ratings for concrete concepts in the ball condition did not differ from the articulatory suppression condition ($BF_{01} = 2.25$, $PP = 13\%$). Finally, there was moderate evidence that difficulty ratings for concrete concepts in the ball condition did not differ from the control condition ($BF_{01} = 3.01$, $PP = 68\%$).

ii) Next, we verified whether the interference effect was particularly strong for tools, and then for food items. We found strong evidence that the interference effect is stronger for tools (vs. the more abstract concepts, i.e. PS) in the ball condition as compared to the control condition ($BF_{10} = 19.9$, $PPs = 99\%$), and extreme evidence that the interference effect is stronger for tools (vs. PS) in the ball condition as compared to the heart beating condition and to the gum condition ($BF_{10} > 100$, $PPs = 100\%$). However, there was moderate evidence that there was no difference with the articulatory suppression condition ($BF_{01} = 7.80$, $PP = 63\%$). Concerning the food, we found inconclusive evidence ($BF_{01} = 1.23$, $PP = 96\%$). We also found extreme evidence that the interference effect was stronger for food items (vs. PS) in the ball condition as compared to the heart beating condition and to the gum condition as compared to the control condition ($BF_{10} > 1000$, $PPs = 100\%$). However, there was moderate evidence that there was no difference with the articulatory suppression condition ($BF_{01} = 4.12$, $PP = 12\%$).

ii) We verified whether within abstract concepts the interference effect was particularly strong for PSTQ (vs. the more abstract concepts, i.e. PS). Within abstract concepts, we found moderate evidence that there was no difference with the control condition ($BF_{01} = 5.9$, $PP = 52\%$). We found extreme evidence that the interference effect was stronger for PSTQ (vs PS) in the ball condition as compared to the heart beating condition and to the gum condition ($BF_{10} > 100$, $PPs = 100\%$). However, there was anecdotal evidence that there was no difference with the articulatory suppression condition ($BF_{01} = 2.23$), although in terms of posterior probabilities it is plausible to assume that the effect was

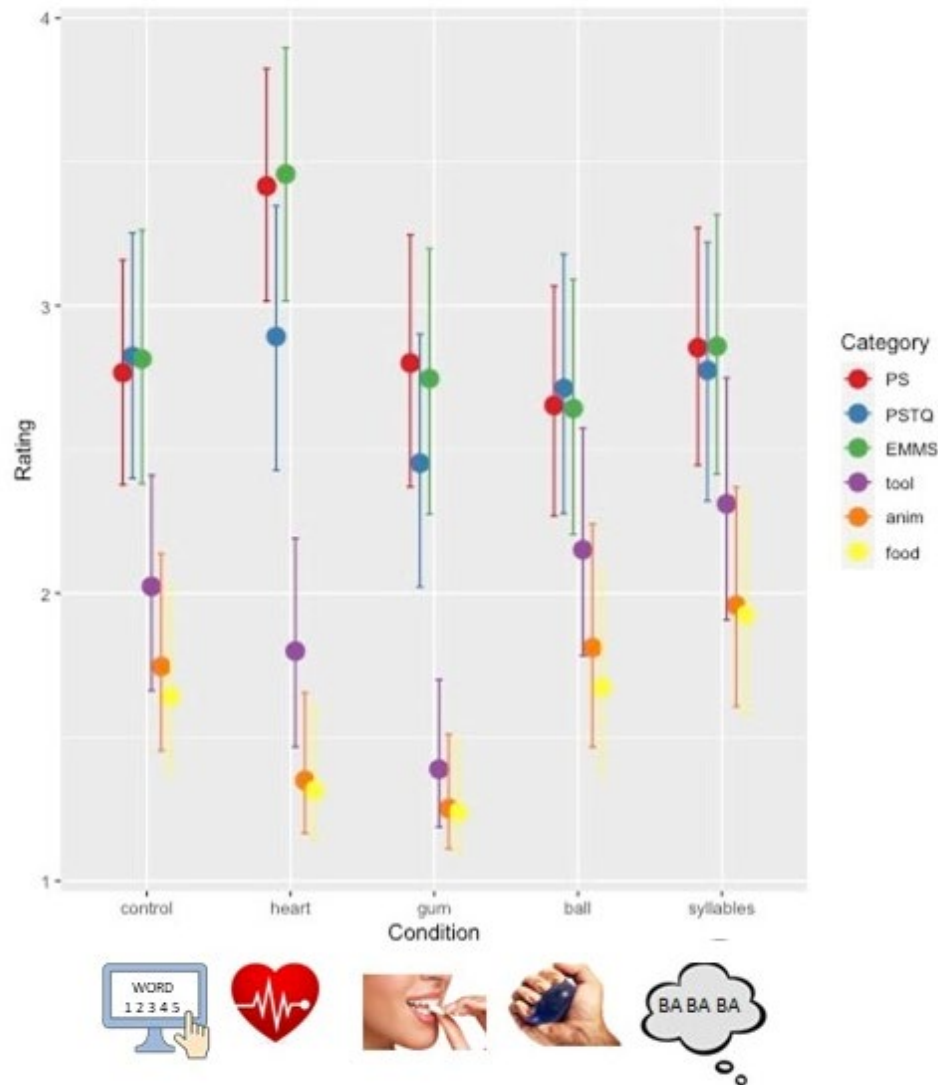
stronger for PSTQ in the ball condition as compared to the articulatory suppression condition (PP = 95%).

Hypothesis 2. i) We predicted that the gum chewing condition would interfere more with abstract concepts than with concrete concepts of animals and tools, determining an increase in difficulty at the increase of the abstractness level. To test this hypothesis, we tested whether the difference between abstract and concrete concepts of animals and tools in the gum condition was different, as compared to the other conditions. When tested against the heart beating condition, we found inconclusive evidence in support of this hypothesis ($BF_{10} = 2.9$), and actually it was more plausible that the difference was in the opposite direction as compared to the predicted one (PP = 0.33%). However, when compared with the control, ball and articulatory suppression conditions, we found extreme evidence in support of our hypothesis ($BF_{10s} > 100$, PPs = 100%). ii) We also predicted that the gum condition would modulate more the food items, either determining a facilitation or an interference. Thus, we compared the difference between the food items and the rest of sub-categories in the gum condition against the same difference in all the other conditions. We found inconclusive evidence for a difference that food items were affected as compared with the control condition ($BF_{10} = 1.1$, PP = 2%). However, we found strong evidence for this hypothesis, when comparing the interference effect on food with the heart condition ($BF_{10} = 84$ because the interference was greater (PP = 100%). When compared to the ball condition, however, we found moderate evidence for this hypothesis ($BF_{10} = 3.89$), but in the opposite direction (PP = .03%), as the interference on food was greater in the ball condition. The same was true in the comparison with the articulatory suppression condition (PP = 0.03%), although in this case the evidence for an effect was extreme ($BF_{10} > 100$).

Hypothesis 3. We predicted i) that the articulatory suppression condition would interfere more with abstract concepts than with concrete concepts, and in particular ii) for the more abstract concepts, i.e. PS). It is clear from a simple visual inspection of the results that hypothesis 3 was not supported by our data (Figure 1), indeed the articulatory suppression condition seems to produce less interference

517 with the abstract concepts, and it was indeed quite similar to the ball condition, as emerged in our
518 analyses related to Hypothesis 2. ii) The same applies to our second sub-hypothesis concerning the
519 more abstract concepts (PS) that did not appear to be judged as more difficult in this condition, as
520 compared to the other experimental conditions (Figure 2).

521 Hypothesis 4. i) We predicted that the heart beating condition would interfere more with abstract
522 concepts than with concrete ones. To test this hypothesis, we tested whether the difference between
523 abstract and concrete concepts was bigger in heart beating condition, as compared to other conditions.
524 We found extreme evidence that the difference in the heart condition was bigger than in all the other
525 conditions, including the control condition ($BF_{10s} > 100$, $PPs = 100\%$). ii) Furthermore, we tested in
526 particular if the effect was bigger for the abstract concepts that involve more the emotional and social
527 dimension. We found extreme evidence for a greater difference between EMSS and PSTQ concepts
528 (PS is the reference level) in the heart beating condition as compared with the ball, the articulatory
529 suppression and the control conditions ($BF_{10s} > 100$, $PPs = 100\%$), moderate evidence for a greater
530 difference between EMSS and PSTQ concepts in the heart condition as compared with the gum
531 condition ($BF_{10} = 7.85$, $PP = 99\%$). ii) Finally, we explored whether the heart beating condition could
532 create more interference with the concepts of animals, because of their animacy. However, even from
533 a simple visual inspection of the results this does not seem to be the case (Figure 2).



534

535 *Fig.2* Interaction plot of ratings mean versus conditions (control, heart, gum, ball, syllables) for the
 536 sub-kinds of abstract (Philosophical and Spiritual concepts, PS; Physical Space Time and Quantity
 537 concepts PSTQ; Emotional, Mental State and Social concepts, EMSS) and concrete concepts (Tools,
 538 Animals, Food). Error bars indicate the 95% credible intervals.

539

540 Exploratory analyses

541 To better interpret how dual-tasks modulated the differences in ratings between the two kinds of
 542 concepts and their sub-kinds, we decided to run further exploratory analyses on our data.

Specifically, we tested whether the difficulty rating for abstract concepts in each condition differed from the rating for abstract concepts in the control condition. The same analysis was conducted for the concrete concepts.

We also tested whether, when compared to the control condition, the difficulty rating in the gum and heart condition was higher for the more abstract concepts, PS and EMSS, than for the most concrete among the abstract concepts, PSTQ.

Finally, we tested whether the perceived difficulty of tools compared to other concrete concepts decreased more in the gum condition than in the control condition.

551

Exploratory analyses results.

Concrete concepts. We found very strong evidence ($BF_{10} = 61.48$) that concrete concepts were judged as less difficult in the gum condition, as compared to the control ($PP = 100\%$). We found only inconclusive evidence ($0.33 < BF_{s10} < 3$) in favor of a difference in the difficulty ratings provided to the concrete concepts between the control condition and the other conditions ($4\% < PP_s < 98\%$).

Abstract concepts. We found strong evidence ($BF_{10} = 15$) that abstract concepts were judged as more difficult in the heart condition, as compared to the control condition ($PP = 100\%$). We found only inconclusive evidence ($0.42 < BF_{s10} < 0.47$) in favor of a difference in the difficulty ratings provided to the abstract concepts between the control condition and the other conditions ($4\% < PP_s < 34\%$).

Differences within abstract concepts. We found extreme evidence ($BF_{s10} > 100$) that PSTQ concepts were considered as less difficult, compared to other abstract concepts, in the heart condition and in the gum conditions as compared to the control condition ($PP_s = 100\%$). We found moderate evidence ($BF_{01} = 6.9$) that PSTQ concepts were not rated differently from other abstract concepts, in the ball condition as compared to the control condition ($PP_s = 32\%$). We found only inconclusive evidence ($BF_{s01} = 1.93$) in favor of the absence of a difference in the difficulty ratings provided to the PSTQ

567 concepts compared to other abstract concepts between the control condition and the other conditions
 568 (PPs = 96%).

569 *Differences within concrete concepts.* We found strong and extreme evidence ($BF_{10} = 94.6$ and BF_{10}
 570 > 100) that tools concepts were considered as more difficult, compared to other concrete concepts, in
 571 the heart condition as compared to the control condition (PP = 99%). We found moderate evidence
 572 ($BF_{s01} > 4$) that tool concepts were not rated differently from other concrete concepts, in the gum (PP
 573 = 13%), in the ball (PP = 86%), and in the syllables condition (PPs = 59%) as compared to the control
 574 condition.

575

576 **Discussion**

577 The results clearly show that the different conditions modulate the ratings of abstract and concrete
 578 concepts, and of sub-kinds of abstract and concrete concepts. In many cases they supported the
 579 hypotheses we had advanced, with some exceptions that we will discuss later. We will summarize
 580 and discuss the implications of our results below.

581 We assume that the increase of difficulty ratings in one condition with respect to the others signal the
 582 presence of an interference. We will focus first on abstract and concrete concepts as a whole, and
 583 then on the sub-kinds of abstract and concrete concepts. Notice that the conditions might differ in
 584 terms of executive demands, but the introduction of a control condition allowed us to have a useful
 585 baseline. While we cannot completely exclude that the comparison between the different conditions
 586 might be impacted by the differences in difficulty between the secondary tasks, we do not think it is
 587 the case. The various conditions differently influenced the ratings on concrete and abstract concepts,
 588 hence we believe that their effect is due to the different dimensions they tackle, and not to the different
 589 level of task difficulty.

590

591 **Abstract and concrete concepts as a whole.**

592 In line with hypotheses 1, 2, and 4, when compared to concrete concepts abstract concepts elicited
593 more interference with the gum chewing and the heart beating condition than with the ball squeezing
594 condition. Results indeed showed that the difference between difficulty ratings in concrete and
595 abstract concepts is larger in the heart beating than in all the other conditions, followed by the gum
596 chewing condition which is larger than in all other conditions with the exception of the heart beating
597 one. This supports the hypothesis that interoceptive experience is crucial for the representation of
598 abstract concepts, and also suggests that processing of more abstract concepts involves the mouth
599 motor system. Exploratory analyses allowed us to determine that the heart beating condition rendered
600 abstract concepts more difficult with respect to all other conditions. The gum chewing condition,
601 instead, rendered concrete concepts easier compared to all the other conditions.

602 As to a possible role of inner speech, our hypothesis that the articulatory suppression interfered more
603 with abstract concepts than with concrete ones was instead not supported.

604 If we focus on concrete concepts, we found that the ball squeezing condition rendered the difference
605 between concrete and abstract concepts smaller compared to the differences in the control, gum, and
606 heartbeat conditions, but not to the articulatory suppression condition. Specifically, the ball squeezing
607 condition rendered concepts more difficult with respect to the gum chewing and to the heart beating
608 conditions, in keeping with our hypothesis that manual activity would interfere more with more
609 concrete concepts. However, there is absence of significant evidence that ball condition renders
610 concrete concepts more difficult than the control and articulatory suppression condition. The
611 difference in difficulty with the control condition is however present when we consider tool concepts,
612 for which manual experience is clearly crucial.

613 In sum, most results confirm our predictions, testifying that abstract concepts are grounded in
614 interoceptive experience and that they evoke the mouth motor system, and that concrete concepts and
615 particularly tools are more grounded in sensorimotor experience and activate the hand motor system.

616 However, with respect to our predictions one result strikes us as novel, and another as unexpected.
 617 The novel result is the pivotal role of interoceptive experience, that strikes us as more crucial than
 618 other dimensions for the representation of abstract concepts.

619 The unexpected result is the scarce modulation of articulatory suppression depending on the
 620 abstractness of stimuli. It is mainly unclear from the results whether articulatory suppression elicited
 621 a selective interference in processing of abstract concepts or instead on both abstract and concrete
 622 ones. In the articulatory suppression condition the disadvantage of abstract over concrete concepts is
 623 slightly larger than in the ball condition, in line with our predictions, but the evidence is inconclusive.
 624 It is therefore possible that the effect of suppression increases the difficulty of all linguistic stimuli,
 625 irrespective of their abstractness level. The result contrasts with recent evidence (Zannino, Fini,
 626 Benassi, Carlesimo, Borghi, under review) in which we found a selective interference of articulatory
 627 suppression on abstract concepts processing, in a task in which we asked participants to judge whether
 628 words were concrete or abstract and we measured response times. It is therefore possible that the
 629 absence of a selective interference due to articulatory suppression is owing to the specific task we
 630 selected, that required participants to explicitly evaluate conceptual difficulty and did not consider
 631 their online performance. Further studies are necessary, to investigate more in depth the role of
 632 articulatory suppression in abstract concepts processing across different tasks.

633 **Sub-kinds of abstract and concrete concepts**

634 PSTQ abstract concepts. As predicted (exploratory hypothesis), we found that the ball squeezing
 635 condition increased difficulty judgments of PSTQ concepts to a larger extent than the heart and gum
 636 conditions, but not than the control condition. Furthermore, as predicted EMSS (together with PS)
 637 differed from PSTQ concepts more in the heart condition compared to all the other conditions. This
 638 result confirms that PSTQ are the most concrete among the abstract concepts, and tap on sensorimotor
 639 (exteroceptive) rather than on interoceptive experience.

640 EMSS abstract concepts. As predicted (directional hypothesis), the heart beating condition interfered
641 in particular with abstract concepts that involve more the emotional and social dimension, i.e., with
642 EMSS, compared with the more concrete PSTQ concepts (but not with PS concepts).

643 Tools concrete concepts. Within concrete concepts, as predicted (directional hypothesis) the ball
644 condition interfered more with judgments on tools when compared with all other conditions except
645 the articulatory suppression one.

646 Food and animals concrete concepts. As predicted (directional hypothesis), compared with the ball
647 squeezing and the suppression condition the gum chewing condition interfered more with abstract
648 concepts than with animal and tool concepts (mouth activation), but also with food ones. Surprisingly,
649 we did not find a clear effect of mouth chewing on food stimuli; instead, concrete concepts were
650 differentiated into the two classical categories of living (food and animals) and nonliving (tools)
651 entities. Interestingly, compared to PS abstract concepts food concepts were considered more difficult
652 in the ball than in the gum and heartbeat condition (but not than in the control and articulatory
653 suppression one), likely because of their graspability. Hence, it appears that food was represented
654 more as graspable, hence more in relation to the hand than to the mouth effector.

655 PS abstract concepts. Our prediction that, because of their higher abstractness level, PS concepts
656 would be mostly interfered in the articulatory suppression condition was not confirmed. This however
657 depended on the fact that, overall, articulatory suppression did not seem to interfere more with
658 abstract concepts than with concrete ones, if not for a slight tendency that requires further studies to
659 be investigated. Interestingly, PS abstract concepts differed from PSTQ ones in interoception, likely
660 because of their higher abstractness level.

661

662 Conclusion

663 The study was aimed to test a general claim and more specific claims deriving from the WAT proposal
 664 (Borghi et al., 2018b, 2019a) and from other proposals on abstract concepts representation. According
 665 to the general claim of the WAT proposal abstract concepts are more characterized than concrete ones
 666 by linguistic experience (see also Dove, 2019, LENS proposal), hence mouth activation, and by inner
 667 grounding and interoceptive experience (see also Connell et al., 2018), and less characterized than
 668 concrete ones by sensorimotor experience related to hand experiences. This general claim was
 669 supported by our results: perceived difficulty of abstract concepts selectively increased when
 670 participants were required to perform a task requiring interoceptive awareness (heart beating
 671 condition). Furthermore, when their mouth active movement was not allowed the processing of
 672 concrete concepts and of the more concrete within abstract concepts, PSTQ, was facilitated,
 673 suggesting the presence of a higher difficulty at the increase of the abstractness level of concepts
 674 (gum chewing condition). Finally, perceived difficulty of concrete concepts, and particularly of tools,
 675 increased when participants had to manipulate an object (ball squeezing condition). Notice that, even
 676 if the instructions we gave did not specify what we intended with “difficulty” of the word, our results
 677 suggest that this was interpreted as difficulty of processing: the words perceived as easier were “dog”
 678 (cane), “grapes” (uva), and “banana” (banana), while the words perceived as more difficult across
 679 conditions were “acceleration” (accelerazione), “enigma” (enigma) and “salvation” (salvezza) (see
 680 supplementary materials).

681 This study was also aimed to test more specific claims concerning the way in which different kinds
 682 of abstract and concrete concepts were represented. Our results demonstrated that abstract concepts
 683 cannot be considered as a whole (Villani et al., 2019), and that different mechanisms underline their
 684 representation. Within abstract concepts, EMSS and PS concepts are more characterized by
 685 interoceptive experience than PSTQ, the more concrete among abstract concepts. Within concrete
 686 concepts, the major differences concerned tools, more grounded in sensorimotor experience (ball
 687 experience) than animals and foods: our results thus confirmed the classic distinction between living
 688 and nonliving entities. Surprisingly, this distinction did not emerge only in the ball squeezing

689 condition, in the direction we expected, but also in the heart beating and articulatory suppression
 690 condition.

691 What diverged from our initial predictions was the pattern elicited by the articulatory suppression
 692 condition, which we expected to provoke selective interference with abstract concepts processing.
 693 Can we conclude that articulatory suppression, typically used to access inner speech (Alderson-Day
 694 & Fernyhough, 2015), has not a selective influence on abstract concepts? Given the discrepant results
 695 found elsewhere with response times (Zannino et al., under review), we are inclined to think that this
 696 condition did not lead to the expected results because of the task, which required an explicit evaluation
 697 and did not have any specific time constraints.

698 Another possibility we can speculate on concerns the mechanisms underlying the mouth motor system
 699 activation. We hypothesized that three mechanisms are at play: a re-enactment of the linguistically
 700 mediated acquisition experience, an inner re-explanation of the word meaning, occurring through
 701 inner speech, and a social metacognitive mechanism, aimed at asking others information to fill our
 702 knowledge gaps. The mechanism for which inner speech is more required is likely the internal re-
 703 explanation of the word meaning. It is possible that this mechanism is less powerful than the others,
 704 at least in the present task. Further studies are needed to investigate this issue.

705 Overall, our study reveals that abstract concepts, compared to concrete ones, are more grounded in
 706 interoceptive and linguistic (mouth motor system) experience, and that abstract concepts are not a
 707 unitary block but that the experiences they rely on widely differ.

708

709 **Acknowledgments and funding information.**

710 We thank Fabrizio Campanile for his collaboration with data collection of the pilot study. We would
 711 like to thank Laura Barca, Chiara Fini, Claudia Mazzuca and Luca Tummolini for comments and
 712 discussions. Funds: Sapienza University of Rome, Progetti medi, “Inner grounding of abstract
 713 concepts: the role of interception and social metacognition” to the last author.

714 **Compliance with ethical standards**

715 **Ethical standards:** The local ethics committee approved the study and it has therefore been
 716 performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki
 717 and its later amendments.

718 **Informed consent:** Informed consent was obtained from all individual participants included in the
 719 study.

720 **Conflict of Interest:** The authors declare that they have no conflict of interest.

721

722 **References**

723

724 Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: development, cognitive functions,
 725 phenomenology, and neurobiology. *Psychological bulletin*, 141(5), 931.

726 Alter, A. L., & Oppenheimer, D. M. (2008). Effects of fluency on psychological distance and mental
 727 construal (or why New York is a large city, but New York is a civilized jungle). *Psychological*
 728 *Science*, 19(2), 161-167.

729 Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory
 730 hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3),
 731 10.1016/j.jml.2012.11.001. doi:10.1016/j.jml.2012.11.001

732 Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of*
 733 *the Royal Society B: Biological Sciences*, 358(1435), 1177-1187.

734 Barsalou, L.W., Dutriaux, L., Scheepers, C. (2018). Moving beyond the distinction between concrete
 735 and abstract concepts. *Phil. Trans. R. Soc. B* **373**, 20170144. (doi:10.1098/rstb.2017.0144)

736 Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M.
 737 (2005). CoLFIS (Corpus e Lessico di Frequenza dell'Italiano Scritto). Available on [Http://Www.](http://www.Istc.Cnr.It/Material/Database)
 738 [Istc. Cnr. It/Material/Database](http://www.Istc.Cnr.It/Material/Database).

- 739 Borghi, A. M., Barca, L., Binkofski, F. & Tummolini, L. (2018a). Abstract concepts, language and
 740 sociality: from acquisition to inner speech. *Philosophical Transactions of the Royal Society of*
 741 *London. Series B, Biological Sciences*, 373 (1752). doi.org/10.1098/rstb.2017.0134
- 742 Borghi, A. M., Barca, L., Binkofski, F. & Tummolini, L. (2018 b). Varieties of abstract concepts:
 743 development, use and representation in the brain. *Philosophical Transactions of the Royal Society*
 744 *of London. Series B, Biological Sciences*, 373 (1752). doi.org/10.1098/rstb.2017.0121
- 745 Borghi, A.M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019a).
 746 Words as social tools: Language, sociality and inner grounding in abstract concepts. *Physics of life*
 747 *reviews*, 29, 120-153. doi: 10.1016/j.plrev.2018.12.001.
- 748 Borghi, A.M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019b).
 749 Words as social tools: Flexibility, situatedness, language and sociality in abstract concepts. Reply
 750 to comments on Words as social tools: Language, sociality and inner grounding in abstract
 751 concepts. *Physics of life reviews*, 29, 178-184. doi: 10.1016/j.plrev.2019.06.004.
- 752 Borghi, A. M. & Binkofski, F. (2014). *Words as social tools: an embodied view on abstract concepts*.
 753 New York; Springer.
- 754 Borghi, A.M., Cimatti, F. (2009). Words as tools and the problem of abstract words meanings. In N.
 755 Taatgen & H. van Rijn (eds.). *Proceedings of the 31st Annual Conference of the Cognitive Science*
 756 *Society* (pp. 2304-2309). Amsterdam: Cognitive Science Society.
- 757 Borghi, A.M., Fini, C., Tummolini, L., (under review). Abstract Concepts and metacognition:
 758 searching for meaning in self and others.
- 759 Borghi, A.M., & Lugli, L. (in prep) Abstract concepts and mouth motor system: evidence with an
 760 interference paradigm.
- 761 Bürkner, P.-C., & Vuorre, M. (2019). Ordinal Regression Models in Psychology: A Tutorial.
 762 *Advances in Methods and Practices in Psychological Science*, 2(1), 77–101.
 763 <https://doi.org/10.1177/2515245918823199>

- 764 Bürkner, P.C. (2017). brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of*
 765 *Statistical Software*. 80(1), 1-28. doi:10.18637/jss.v080.i01
- 766 Bürkner, P.C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R*
 767 *Journal*. 10(1), 395–411. doi:10.32614/RJ-2018-017
- 768 Champely, S. (2018). pwr: Basic functions for power analysis (ver.1.2-2.). Retrieved from
 769 <https://CRAN.Rproject.org/package=pwr>.
- 770 Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L.
 771 Erlbaum Associates.
- 772 Connell, L., & Lynott, D. (2012). When does perception facilitate or interfere with conceptual
 773 processing? The effect of attentional modulation. *Frontiers in Psychology*, 3, 474.
- 774 Connell, L., Lynott, D., & Banks, B. (2018). Interoception: the forgotten modality in perceptual
 775 grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society of*
 776 *London. Series B, Biological Sciences*, 373(1752). <https://doi.org/10.1098/rstb.2017.0143>
- 777 Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R package version
 778 2019.4-25. <http://www.cran.r-project.org/package=ordinal/>.
- 779 Crutch, S.J., Troche, J., Reilly, J., Ridgway, G.R. (2013). Abstract conceptual feature ratings: the role
 780 of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual
 781 knowledge. *Front Human Neurosci*; 7:186. <https://doi.org/10.3389/fnhum.2013.00186>.
- 782 Della Rosa, Pasquale A., Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract
 783 concrete dichotomy: mode of acquisition, concreteness, imageability, familiarity, age of
 784 acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior*
 785 *Research Methods*, 42(4), 1042–1048. <https://doi.org/10.3758/BRM.42.4.1042>
- 786 Desai, R. H., Reilly, M., & van Dam, W. (2018). The multifaceted abstract brain. *Philosophical*
 787 *Transactions of the Royal Society of London. Series B, Biological Sciences*, 373(1752).
 788 <https://doi.org/10.1098/rstb.2017.0122>

- 789 Dienes, Z. (2008). *Understanding Psychology as a Science: An Introduction to Scientific and*
790 *Statistical Inference*. Palgrave-Macmillan.
- 791 Dove, G., (2019). More than a scaffold: Language is a neuroenhancement. *Cogn Neuropsychol*, 4:1-
792 24. doi: 10.1080/02643294.2019.1637338
- 793 Dreyer, F. R., & Pulvermüller, F. (2018). Abstract semantics in the motor system?—An event-related
794 fMRI study on passive reading of semantic word categories carrying abstract emotional and mental
795 meaning. *Cortex*, 100, 52-70.
- 796 Falandays, J.B., & Spivey, M. J. (2019). Abstract meanings may be more dynamic, due to their
797 sociality: Comment on" Words as social tools " by Borghi et al. *Physics of life reviews*, 29, 175-
798 177.
- 799 Fingerhut, J., Prinz, J.J., (2018). Grounding evaluative concepts. *Phil. Trans. R. Soc. B* 373,
800 20170142. (doi:10.1098/rstb.2017.0142)
- 801 Fischer, M.H., Shaki, S., (2018). Number concepts: abstract and embodied. *Philos Trans R Soc Lond*
802 *B, Biol Sci* 2018; 373 (1752). <https://doi.org/10.1098/rstb.2017.0125>.
- 803 Forde & Humphreys, (2005). *Category specificity in brain and mind*. Hove, East Sussex: Psychology
804 Press; New York: Taylor & Francis.
- 805 Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your
806 own heart: distinguishing interoceptive accuracy from interoceptive awareness. *Biological*
807 *psychology*, 104, 65-74.
- 808 Ghio, M., Vaghi, M. M. S., & Tettamanti, M. (2013). Fine-grained semantic categorization across the
809 abstract and concrete domains. *PloS one*, 8(6).
- 810 Glenberg, A. M. (2019). Turning social tools into tools for action. Comment on" Words as social
811 tools " by Borghi et al. *Physics of life reviews*, 29, 172-174.
- 812 Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate*
813 *behavioral research*, 26(3), 499-510

- 814 Hoffman, P. (2016). The meaning of 'life' and other abstract words: Insights from neuropsychology.
815 *Journal of Neuropsychology*, 10, 317-343.
- 816 Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of
817 semantic ambiguity based on variability in the contextual usage of words. *Behavior Research*
818 *Methods*, 45, 718–730.
- 819 Jeffreys, H. (1961). *Theory of Probability*. 3rd Edition, Clarendon Press, Oxford.
- 820 Keil, F.C. (1989). *Concepts, kinds and cognitive development*, MIT Press, London.
- 821 Kousta, S. T., Vigliocco, G., Vinson, D., Andrews, M., & Del Campo, E. (2011). The motor and pre-
822 motor cortex. *Neuron*, 41, 301–307. doi: 10.1016/s0896-6273(03)00838-9
- 823 Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. Age-of-acquisition ratings for 30,000
824 English words. *Behav Res* 44, 978–990 (2012). <https://doi.org/10.3758/s13428-012-0210-4>
- 825 Lee, M. D., & Wagenmakers, E.J. (2013). *Bayesian cognitive modeling: A practical course*.
826 Cambridge University Press.
- 827 Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could
828 possibly go wrong? *Journal of Experimental Social Psychology*, 79, 328-348.
829 <http://dx.doi.org/10.1016/j.jesp.2018.08.009>
- 830 Lugli, L., & Borghi, A.M. (2017). Chewing a gum interferes with abstract concepts' processing. 20th
831 *Conference of the European Society for Cognitive Psychology* (ESCOP 2017). Postdam, Germany,
832 3-6 September 2017.
- 833 Lupyan, G. (2019). Language as a source of abstract concepts: Comment on "Words as social tools:
834 Language, sociality and inner grounding in abstract concepts" by Anna M. Borghi et al. *Physics*
835 *of Life Reviews*, 29, 154-156. doi: 10.1016/j.plrev.2019.05.001. Epub 2019 May 13.
- 836 Martin, A. (2007). The representation of object concepts in the brain. *Annu Rev Psychol*, 58:25-45.
- 837 Murphy, G. (2002). *The Big Book of Concepts*, Cambridge, MA: MIT Press.
- 838 Newcombe, P.I., Campbell, C., Siakaluk, P.D., & Pexman, P.M. (2012). Effects of emotional and
839 sensorimotor knowledge in semantic processing of concrete and abstract nouns. *Frontiers in*
840 *human neuroscience*, 6, 275.

- 841 Nezlek, John B.(2011) Multilevel modeling for social and personality psychology. SAGE
842 Publications Ltd.
- 843 Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.
- 844 Prinz, J. J. (2014). *Beyond human nature: How culture and experience shape the human mind*. WW
845 Norton & Company.
- 846 Pulvermüller, F. (2018). The case of CAUSE: neurobiological mechanisms for grounding an abstract
847 concept. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752),
848 20170129.
- 849 Rumiat, R.I. & Foroni, F. (2016). We are what we eat: How food is represented in our mind/brain.
850 *Psychon Bull Rev*, 23: 1043. <https://doi.org/10.3758/s13423-015-0908-2>
- 851 Schandry, R. (1981). Heartbeat perception and emotional experience. *Psychophysiology*, 18(4), 483–
852 488.
- 853 Shea, N. (2018). Metacognition and abstract concepts. *Philosophical Transactions of the Royal*
854 *Society of London. Series B, Biological Sciences*, 373(1752).
855 <https://doi.org/10.1098/rstb.2017.0133>
- 856 Topolinski, S., & Strack, F. (2009). Motormouth: mere exposure depends on stimulus-specific motor
857 simulations. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35(2), 423–
858 433. <https://doi.org/10.1037/a0014504>
- 859 Topolinski, S., Lindner, S., & Freudenberg, A. (2014). Popcorn in the cinema: Oral interference
860 sabotages advertising effects. *Journal of Consumer Psychology*, 24(2), 169–176.
- 861 Vigliocco, G., Kousta, S.T., Della Rosa, P.A., Vinson, D. P., Tettamanti, M., Devlin, J.T., & Cappa,
862 S. F. (2014). The neural representation of abstract words: the role of emotion. *Cerebral Cortex*,
863 24(7), 1767-1777.
- 864 Villani C., Lugli L., Liuzza MT., Borghi AM. (2019). Varieties of abstract concepts and their multiple
865 dimensions, *Language and Cognition* 11(3), 403-430. doi:10.1017/langcog.2019.23
- 866 Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic*
867 *bulletin & review*, 14(5), 779-804.

- 868 Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, 107(3),
869 829–27 853.
- 870 Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable
871 information criterion in singular learning theory. *Journal of Machine Learning Research*, 11(Dec),
872 3571-3594.
- 873 Wauters LN, Tellings AE, Van Bon WH, Van Haaften AW. (2003). Mode of acquisition of word
874 meanings: the viability of a theoretical construct. *Applied Psycholinguistics* 24(03):385406 DOI
875 10.1017/S0142716403000201.
- 876 Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing
877 fluency: Implications for evaluative judgment. In J. Musch, K.C. Klauer (ed.) *The psychology of*
878 *evaluation: Affective processes in cognition and emotion*, 189, 217.
- 879 Zannino, GD., Fini, C., Benassi, M., Carlesimo G.A., Borghi, A.M. (under review) *Articulatory*
880 *suppression and processing of abstract concepts: the role of inner speech*.
- 881

882 **Figure captions:**

883

884 **Fig. 1:** Interaction plot of ratings mean versus conditions (control, heart, gum, ball, syllables) for
885 abstract and concrete concepts. Error bars indicate the 95% credible intervals.

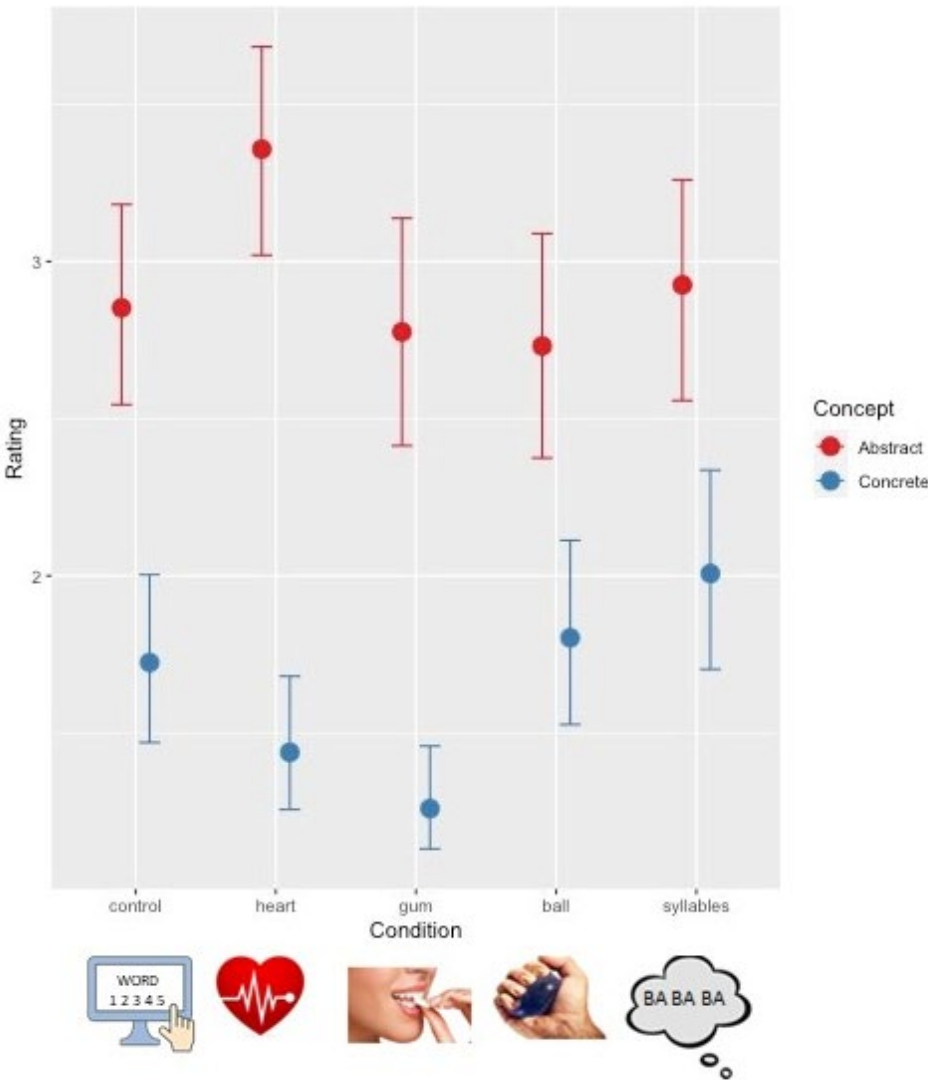
886

887 **Fig. 2:** Interaction plot of ratings mean versus conditions (control, heart, gum, ball, syllables) for the
888 sub-kinds of abstract (Philosophical and Spiritual concepts, PS; Physical Space Time and Quantity
889 concepts PSTQ; Emotional, Mental State and Social concepts, EMSS) and concrete concepts (Tools,
890 Animals, Food). Error bars indicate the 95% credible intervals.

891

892 **Fig. 1**

893

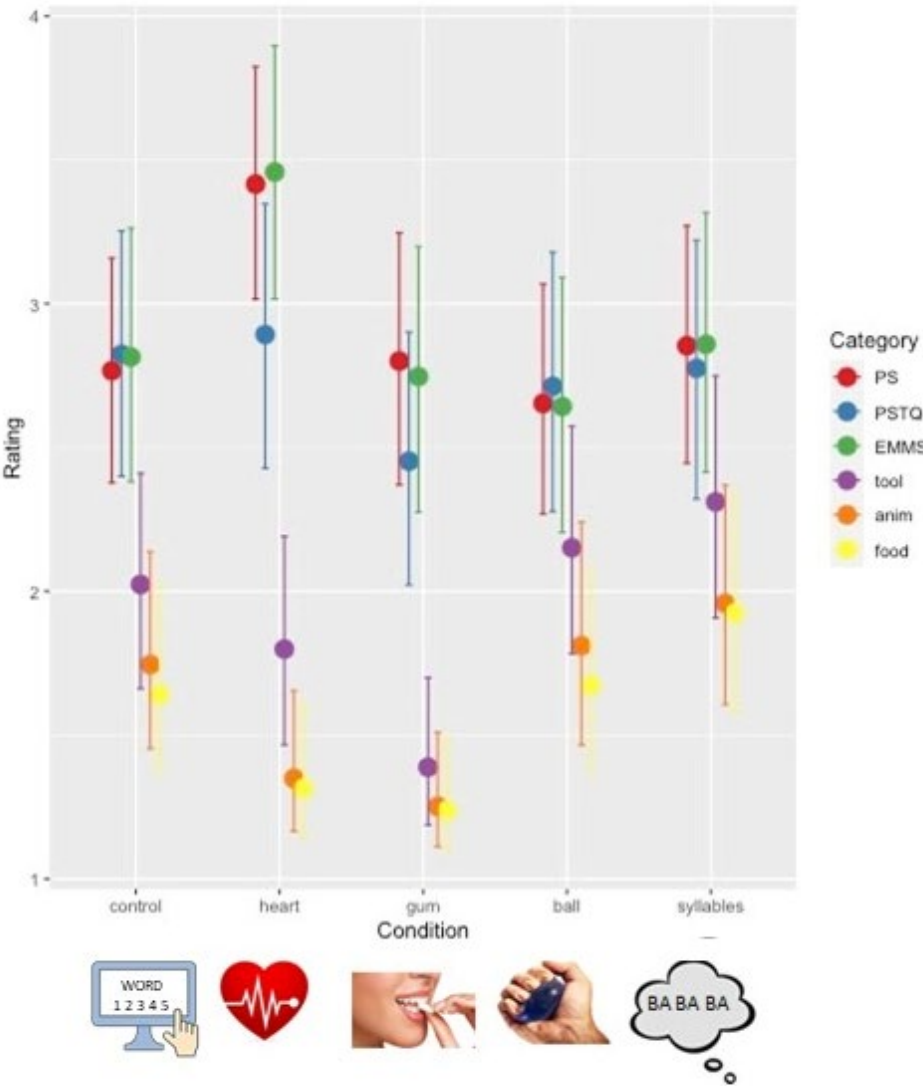


894

895

896 Fig.2

897



898

899

900