7

Different Effects of Word Concreteness in a Recall and Recognition Task

Dunja Mićunović
FMK Lab—Laboratory for Psychological Research,
Department of Psychology, Faculty of Media and Communications
dunja.micunovic.20211057@fmk.edu.rs

Anja Lujić
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia

Senka Vasović
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia

Marko Janković
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia

Milica Popović Stijačić
Department of Psychology, FMK Lab—Laboratory for Psychological Research,
Faculty of Media and Communications, Serbia
Laboratory for Experimental Psychology, Faculty of Philosophy,
University of Novi Sad, Serbia

Abstract

Based on Dual Coding Theory, Depth of Processing Theory, and Perceptual Symbols Systems Theory, we examined the relationship between the concreteness effect and depth of processing in memory tasks. We also manipulated the number of sensory modalities associated with concrete words to test the peg hypothesis, which suggests that additional sensory cues enhance recall. The participants were randomly assigned to one of three semantic rating tasks—concreteness, context availability, or vividness—and rated words accordingly. The words were categorised as abstract, few-modalities, or many-modalities based on the number of sensory experiences they evoke. After completing the rating task, all the participants proceeded to a mental rotation task, followed by random *assignment to one of three previously unannounced memory tasks: free recall, cued recall, or recognition. As expected, the vividness group achieved the highest retrieval across all tasks, while the other groups showed no clear pattern. Notably, abstract words were recalled most accurately in the cued recall task, while many-modality concrete words led to the highest recognition accuracy. These findings are discussed using the Dynamic Visual Noise Paradigm, as the mental rotation task may have interfered with the visualisation processes involved in consolidating concrete word representations. Our results do not support the peg hypothesis or the assumptions of Perceptual Symbols Systems Theory since no recognition difference was found between few- and many-modality concrete words.*

Keywords: *Dual Coding Theory, Depth of Processing Theory, Embodiment Theory, incidental learning, concreteness effect, memory tasks*

Introduction

The concreteness effect is one of the most extensively studied phenomena in cognitive psychology, yet a definitive explanation for this effect remains elusive. It refers to the advantage that concrete words have over abstract words in word-processing tasks—namely, shorter processing latencies for concrete words and higher accuracy in recall and recognition (Doest & Semin, 2005; Hamilton & Rajaram, 2001; Kounios & Holcomb, 1994; Paivio, 1991; Paivio et al., 1994; Popović Stijačić & Filipović Đurđević, 2015, 2022; Romani et al., 2008; Schwanenflugel & Stowe, 1989; Schwanenflugel et al., 1992; Vandendaele & Grainger, 2022). The two most prominent theoretical accounts are the *Dual Coding Theory* (DCT; Paivio, 1991) and the *Context Availability Theory* (Schwanenflugel et al., 1992).

Theoretical Explanations of Concreteness Effect

The DCT (Paivio, 1991) postulates that concrete words have a memory advantage because they are encoded through additional sensory information besides the symbolic code they share with abstract words. These analogue codes serve as extra cues (i.e., pegs) during retrieval. This assumption is known as the *peg hypothesis* (Paivio, 1991, 1994). The theory emphasises the role of visual and other sensory components in word processing and recall, suggesting that concrete words are more easily remembered than abstract words due to their specificity and ease of visualisation. Over the past several decades, the importance of sensory-motor information related to word meaning has been highlighted by the Perceptual Symbol Systems Theory—one of the main approaches within embodiment theories—which postulates that our knowledge is grounded in the sensory-motor system (Barsalou, 1999, 2004). The Dual Coding Theory and the Perceptual Symbol Systems Theory are complementary, as the latter decomposes sensory-motor experience and distinguishes between words not based on general concreteness but through fine-grained measures such as the number of sensory modalities (NoM) through which a word's meaning can be experienced (Filipović Đurđević et al., 2016; Popović Stijačić & Filipović Đurđević, 2015, 2022).

Previous research on the Serbian language (Popović Stijačić & Filipović Đurđević, 2015, 2022) has tested the extension of the peg hypothesis in relation to the Dual Coding Theory (DCT) and Perceptual Symbol Systems Theory, examining both free and cued recall (Paivio, 1991; Paivio et al., 1994). According to this hypothesis, analogue codes related to concrete words serve as additional cues during retrieval. Merging the ideas of both DCT and Perceptual Symbol Systems Theory, the authors hypothesised that words whose meanings can be experienced through more sensory modalities would show better recall accuracy. As predicted, reproduction accuracy increased with the number of sensory modalities through which the meaning of a word could be experienced. More specifically, the lowest reproduction accuracy was observed for abstract words (e.g., "truth"), which cannot be experienced through the senses, followed by concrete words whose meanings can be experienced through one or two senses (e.g., "moon"). The highest reproduction accuracy was recorded for concrete words whose meanings can be experienced through more than two senses (e.g., "fire"). In other words, a greater number of sensory modalities through which a word's meaning can be experienced serves as a greater number of "pegs" or additional cues during retrieval. Still, this effect was present only in the cued recall task, not in the free recall task.

Contrary to DCT and Perceptual Symbol Systems, the Context Availability Theory (Schwanenflugel & Stowe, 1989) posits that the advantage of concrete words stems from their richer contextual information, which is more readily available than that of abstract words during word processing and recall. Namely, Schwanenfluegel et al. (1992) found that the concreteness effect is the consequence of strategic imagery usage during encoding and retrieval, and that depends on individual differences. Thus, the individuals who reported using imagery during encoding and retrieval had greater recall accuracy of the concrete words.

Depth of Processing and Recall

Concerning recall accuracy in general, the encoding phase of words plays a crucial role. With this in mind, we draw on the assumptions of the Depth of Processing (DoP) Theory (Craik & Lockhart, 1972), according to which words are processed at varying levels of depth, from shallow sensory processing to deeper semantic processing. According to DoP, deeper processing involves attentional mechanisms oriented towards the meaning of stimuli, resulting in more durable memory traces. The deeper the processing, the better the recall or reproduction of the word.

By integrating multiple theoretical perspectives, including Dual Coding Theory (DCT), Perceptual Symbol Systems, Context Availability, and the Depth of Processing (DoP) framework, we aim to understand better the role of incidental learning in retrieving concrete and abstract words. Specifically, by directing attention to different semantic properties of words, we hypothesised that it is possible to manipulate the depth of processing during incidental encoding, a method similar to that used by West and Holcomb (2000).

For example, we examined whether the concreteness effect could be enhanced if participants had previously rated words based on concreteness or vividness of mental imagery rather than context availability. Although all three types of ratings are semantic, both concreteness and vividness of mental imagery are more likely to engage additional sensory memory cues. Consequently, these assessments may require greater cognitive effort and time, leading to more thorough processing and improved recall.

West and Holcomb (2000) applied a similar procedure in a sentence verification task. They found that the concreteness effect was strongest when participants made semantic judgments about the sentence, somewhat reduced when they performed imagery-based judgments, and absent in a shallow task (e.g., "Is the letter *n* present in the target word?").

Study Aims and Hypotheses

Our study had two primary aims. (1) To investigate the concreteness effect across different memory tasks (free recall, cued recall, and recognition), using a definition of concreteness based on the number of sensory modalities through which a concept can be experienced. This variation allowed us to test the peg hypothesis more directly by comparing two groups of concrete words differing in their number of potential sensory "pegs": words associated with one or two modalities ("few-modality concrete") versus those linked to three or more modalities ("many-modality concrete").

(2) To examine how the different aspects of semantic processing influence memory performance as an extension of the Depth of Processing (DoP) Theory. Specifically, we investigated whether directing attention to various semantic properties of words during encoding, namely concreteness, context availability, and vividness of mental imagery, would enhance the recall of abstract and concrete words differently. Depth of processing was additionally explored by the time the participants spent rating words during the encoding, assuming that more time reflects deeper processing.

We hypothesised that the main effect of the memory task would be such that recognition would yield the highest retrieval accuracy, followed by cued recall, with the lowest accuracy expected in free recall.

Additionally, based on previous research, we expected the main effect of the number of sensory modalities associated with word meanings. Specifically, we predicted the highest accuracy for concrete words representing concepts that can be experienced through many modalities, somewhat lower accuracy for concrete words representing concepts experienced through fewer modalities, and the lowest accuracy for abstract words.

We also examined how different types of semantic rating tasks—used during the encoding phase—might influence memory performance by manipulating the depth of processing. We expected vividness ratings to be the most cognitively demanding, as they require participants to generate mental images of word meanings and evaluate their vividness. The concreteness ratings were assumed to be less demanding, as they involve evaluating whether the concept can be experienced through the senses without necessarily generating a mental image. The context availability ratings, while similarly effortful, were thought to involve a different attentional focus—directed more towards referential and situational knowledge than sensory experience.

Based on this, we predicted that the concreteness effect would be more substantial when participants rated words for vividness or concreteness and weaker when they rated context availability due to differences in the nature and depth of semantic processing.

To test these assumptions, we conducted an experiment in which the participants were randomly assigned to one of three semantic assessment groups: concreteness, context availability, or vividness. The participants rated a list of words in each group according to the assigned dimension. The words varied by the number of sensory modalities through which their meanings can be experienced and were categorised as abstract, few-modalities, or many-modalities words. After completing the rating task, the participants were again randomly assigned to one of three unannounced memory tasks: free recall, cued recall, or recognition. This design allowed us to examine the interaction between the type of semantic processing during encoding and the type of memory task during retrieval.

Method

Participants

A total of 273 psychology students from the Faculty of Media and Communications at Singidunum University in Belgrade participated in the study (85.7% female), with a mean age of 22.78 **(**SD **=** 5.75). The participants were randomly assigned to one of three incidental learning groups, with group sizes as follows: n (G1) = 47, n (G2) = 107, and n (G3) = 46. Table 1 presents the distribution of the participants across the assessment and retrieval groups.

Table 1
Number of participants per assessment and retrieval groups

		Retrieval task			Total
		Free recall	Cued recall	Recognition	
Assessment group	Concreteness	23	23	26	72
	Context availability	83	24	24	131
	Vividness	23	24	23	70
Total		129	71	73	273

Stimuli

The list comprised 45 target Serbian nouns selected from a normative study (Popović Stijačić, 2021), including 15 abstract nouns (e.g., *custom*) and 30 concrete nouns. The concrete nouns were divided into two groups: 15 nouns referred to concepts that can be experienced through one or two sensory modalities (*few-modality words*, e.g., *jaw*), and 15 referred to concepts that can be experienced through more than two modalities (*many-modality words*, e.g., *bell*), thereby operationalising the Number of Modalities (NoM) factor. All three noun groups were matched on a wide range of psycholinguistic variables, including objective word frequency (Kostić, 1999), familiarity (subjective frequency), context availability, emotional valence, arousal, and age of acquisition (Popović Stijačić, 2021). The two concrete noun groups were also matched for concreteness and imageability. This careful matching ensured that any observed differences in retrieval accuracy could be specifically attributed to the perceptual richness of the words, that is, the number of sensory modalities (NoM) through which a concept can be experienced. Four filler words were added at both the beginning and end of the list to control for primacy and recency effects (Murdock, 1962; Glanzer & Cunitz, 1966). Descriptive statistics for the psycholinguistic variables used to match the three NoM word groups are presented in Table 2.

Table 2
Descriptive statistics and F test for the target words by number of modalities factor

		N	Mean	SD	*F*-test	*p*-value
logF	Abstract	15	4.113	1.363		
	Few-modalities	15	3.476	1.445	2.094	0.136
	Many-modalities	15	2.957	1.806		
Length	Abstract	15	6.067	1.335		
	Few-modalities	15	6.400	1.404	0.213	0.809
	Many-modalities	15	6.333	1.676		
Imageability	Abstract	15	3.866	1.211		
	Few-modalities	15	6.317	0.452	50.512	0.000
	Many-modalities	15	6.362	0.381		
Context availability	Abstract	15	5.508	0.527		
	Few-modalities	15	5.448	0.422	0.258	0.774
	Many-modalities	15	5.591	0.668		
EV	Abstract	15	4.675	1.086		
	Few-modalities	15	4.016	0.754	2.301	0.113
	Many-modalities	15	4.217	0.692		
Familiarity	Abstract	15	5.405	0.994		
	Few-modalities	15	5.167	0.629	0.454	0.638
	Many-modalities	15	5.403	0.695		
Concreteness	Abstract	15	3.077	0.844		
	Few-modalities	15	6.079	0.374	124.408	0.000
	Many-modalities	15	6.044	0.471		

Note. logF = logarithm of objective frequency; length = number of letters; EV = emotional valence; N = number of words per group; SD = standard deviation.

In the cued recall task, an additional set of 45 cue words was selected to accompany the target words during retrieval. These cues were either associatively related (based on experiential relatedness, e.g., *weight—scale*) or semantically related (belonging to the same conceptual category, e.g., *lie—truth*) to the target words. The complete list of target—cue pairs is provided in the Appendix (Table A). The same cue words were also used as distractors in the subsequent recognition task.

Procedure

The entire experiment was administered online using the SoSci Survey (Leiner, 2023). Figure 1 illustrates the organisation and timeline of the experiment. Initially, the participants provided informed consent. Since the experiment included unannounced memory tasks, the whole purpose of the study was disclosed to the participants during the debriefing, which took place after they had completed the experiment.

Incidental learning phase:

The participants were randomly assigned to one of three groups based on the type of incidental learning tasks. These tasks involved estimating either word concreteness, context availability, or the vividness of mental images evoked by the words. The words were presented simultaneously as a list, with rating scales placed next to each word. Above the list, the participants saw instructions corresponding to their assigned task. For concreteness, the participants rated the extent to which a word denotes something that can be experienced through the senses. For context availability, the participants rated how easily a meaningful context could be evoked for the word. To assess the vividness of mental images, the participants were asked to imagine the word and rate the vividness of the mental image that was produced. No mention was made of the upcoming recall tasks.

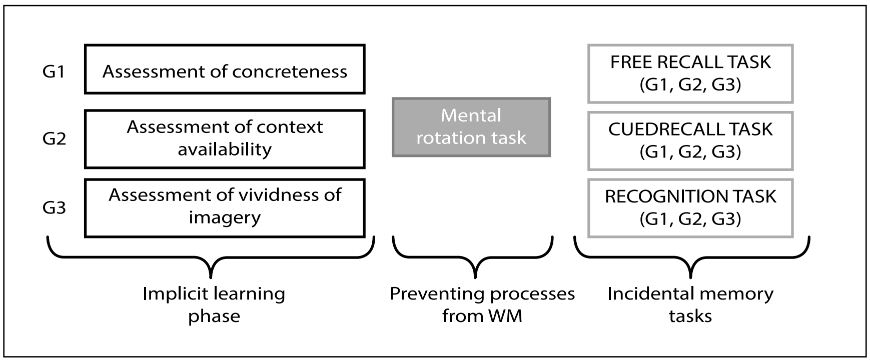
Mental rotation task:

After completing the incidental learning phase, the participants performed a mental rotation task. This task was designed to increase retrieval difficulty and minimise the contribution of working memory. It included 96 experimental trials, preceded by eight practice trials. Two geometric figures were presented on the screen, with the figure on the right rotated at one of four angles (0°, 50°, 100°, or 150°). The participants had to mentally rotate the figure and judge whether it matched the figure on the left. If the figures matched, they pressed the "V" key; if not, they pressed "M". The average time for this task was 6.1 minutes (SD = 3.23 minutes).

Incidental retrieval phase:

The participants were then randomly assigned to one of three retrieval tasks. In free recall, the participants were asked to recall as many words as possible from the incidental learning phase (i.e., the rating task) within a five-minute time limit. In cued recall, the participants were given associatively or semantically related cues to the target words to help them recall the previously estimated words. Reproduction in this task was also limited to five minutes.

Figure 1
Research flow of the experiment



In the recognition task, the participants were presented with a random sequence of previously seen ("old") and unseen ("new") words. Each trial began with a fixation cross (500ms), and words remained on the screen until the participant responded by pressing "V" for old words and "M" for new words.

Variables

Although all three memory tasks had the same stimuli and design, we analysed the answers from free and cued recall separately from the recognition answers.

We employed a 3x3x3 mixed-factor design, with the incidental learning group and memory task as between-participants and within-stimuli factors and the Number of Modalities (NoM) as a between-stimuli and within-participants factor. The incidental learning group consisted of three levels: assessment of concreteness, context availability, and vividness of mental imagery. The memory task consisted of two levels: free recall and cued recall. Finally, NoM had three levels: abstract, few-modality, and many-modality nouns.

The dependent variable was the proportion of correct reproductions in free and cued recall, averaged across the participants. In the recognition task analysis, we used the proportion of correctly recognised targets (hits) corrected for false alarms.

Additionally, we measured the time spent in the encoding phase (i.e., during the assessment of the words) to test the assumptions of the DoP theory. We assumed that more time spent on word assessment reflected deeper processing of the stimuli.

Statistical Analysis

Since we manipulated three factors, we conducted a three-way mixed-effects ANOVA for the by-participant analysis of the recall data, and a two-way mixed-effects ANOVA was applied for the recognition data. We performed a one-way ANOVA for independent samples to test the DoP assumptions regarding the time spent in the incidental learning phase. Data were analysed using SPSS for Windows, version 20 (IBM Corp., 2011).

Results and Discussion

Two participants from the recognition task were excluded from further analysis, as their accuracy, after correcting hits for false alarms, was below zero, indicating random responses.

First, we examined the relationship between encoding time and assessment type to investigate the assumption that vividness ratings would necessitate the most extensive word processing. We conducted a one-way ANOVA with assessment type as the independent variable and encoding time (i.e., the time spent on ratings during the incidental learning phase) as the dependent variable. A significant main effect of assessment type was found: $F(2, 268) = 13.70$, $p < .001$, $\eta^2 = .093$. Table 3 presents the mean encoding times across different assessment and memory tasks. Although vividness ratings required the most time compared to the other assessment groups, post hoc tests with Bonferroni correction revealed significant differences between vividness and context availability ratings (Mdiff = 126.89, $p < .001$), as well as between concreteness and context availability ratings (Mdiff = 95.81, $p < .001$). These results are not entirely consistent with the hypothesis that vividness ratings would be the most demanding compared to both concreteness and context availability ratings.

Table 3
Mean encoding time across assessment group and memory tasks

Assessment group		Mean	SD	N
Concreteness	Free recall	297.35	160.49	23
	Cued recall	292.43	136.68	23
	Recognition	339.24	191.51	25
	Total	310.51	164.32	71
Context availability	Free recall	188.75	76.90	83
	Cued recall	252.08	165.37	24
	Recognition	269.35	138.97	23
	Total	214.70	114.69	130
Vividness	Free recall	350.13	275.17	23
	Cued recall	272.17	131.05	24
	Recognition	405.48	347.59	23
	Total	341.59	267.30	70

Analysis of Recall Data

The three-way mixed-effects ANOVA revealed a significant main effect of the recall task: $F(1, 194) = 16.053$, $p < .001$, $\eta^2 = .076$. As hypothesised, retrieval accuracy was higher in the cued recall condition. Table 4 presents retrieval accuracy across the different memory tasks.

Table 4
Mean retrieval accuracy in recall tasks

	Mean	SE	N
Free recall	.193	.12	129
Cued recall	.264	.13	71

Note. SE = standard error.

Furthermore, a main effect of the assessment group was observed: $F(2, 194) = 4.813$, $p < .01$, $\eta^2 = .047$. Table 5 presents the proportion of retrieval accuracy across the assessment groups. Consistent with our assumptions, the highest recall was achieved in the vividness group, while the concreteness and context availability groups showed similar levels of accuracy (*M*diff = –.01, *p* = 1.00). Post hoc analysis with the Bonferroni correction revealed that the vividness group had significantly better accuracy than both the concreteness group (*M*diff = .066, $p < .01$) and the context availability group (*M*diff = .055, $p < .01$).

Table 5
Mean retrieval accuracy across the assessment groups

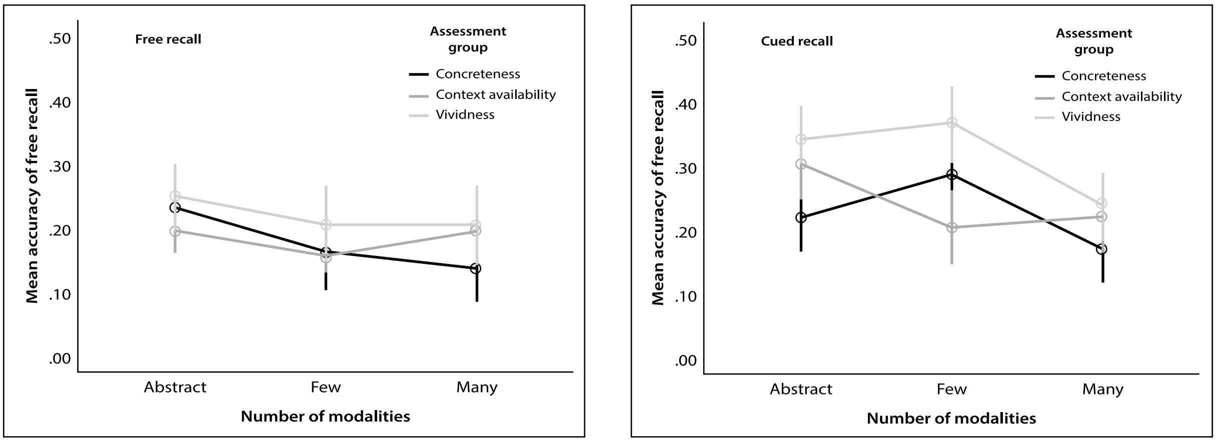
	Mean	SE	N
Concreteness	.203	.017	46
Context availability	.214	.013	107
Vividness	.269	.016	47

Note. SE = standard error.

In contrast to previous research (Paivio et al., 1994; Popović Stijačić & Filipović Đurđević, 2015, 2022), the highest recall accuracy was obtained for abstract nouns (*M* = .259 ± .14), then for few-modality nouns (*M* = 20.93 ± .16), and eventually for many-modalities (*M* = 19.6 ± .14), meaning reversed NoM effect: $F(2, 388) = 16.13$, $p < .001$, $\eta^2 = .076$. The Post Hoc analysis with Bonferroni correction revealed that abstract words significantly differed from few- (*M*diff = .282, *p* = .044) and many-modalities nouns (*M*diff = .626, *p* < .001), where few-modalities also significantly differ from many-modalities nouns (*M*diff = .345, *p* = .005). We also recorded the NoM and assessment task: $F(4, 388) = 4.85$, $p < .001$, $\eta^2 = .048$, where a similar pattern of recall, a reversed NoM effect, was observed for the vividness and concreteness assessment groups, and a different pattern for the context availability group. The interaction of NoM and recall task reached statistical significance: $F(4, 388) = 6.35$, $p = .002$, $\eta^2 = .032$: a reversed NoM effect was present in the cued but not in the free recall task. A significant triple interaction was also obtained: $F(4,388) = 4.495$, $p = .001$, $\eta^2 = .044$ (Figure 2). Post Hoc tests with Bonferroni correction showed that none of the differences, regardless of the assessment group, were significant in free recall. In cued recall, the abstract words in the vividness assessment group had a higher recall than the many-modalities words (*M*diff = .108, *p* = .020). Furthermore, in the same task, two groups of concrete words differed, resulting in the lowest recall accuracy for many-modalities groups (*M*diff = .131, *p* < .001). In the concreteness assessment group, only two groups of concrete words differed, again in the same direction as in the vividness group (*M*diff = .11, *p* = .003), and the abstract words did not differ from either of the concrete word groups. There were no differences in the cued recall accuracy between word groups in the context availability group.

Figure 2

Interaction of NoM and assessment group in free recall (left panel) and cued recall (right panel)



Note. The vertical lines denote 95% confidence intervals of the group means.

The depth of processing assumptions concerning recall accuracy was partially confirmed, as the vividness assessment group achieved better reproduction than the other two assessment groups, and the concreteness and context availability groups had a similar recall rate. However, comparing the assessment time, the vividness and concreteness groups did not significantly differ, indicating that the processing time cannot entirely explain the depth of processing and that some other processes are involved, such as intentional mental visualisation during vividness assessment (the visualisation was part of the instruction).

The reversed NoM effect implies that the participants retrieved abstract words with less effort, although by definition, they are relying only on symbolic code and associative semantic networks. Further, the fine grain division of concrete words showed that, contrary to previous findings (Popović Stijačić & Filipović Đurđević, 2015; 2022), words denoting concepts with richer perceptual experience (many-modalities nouns) had the lowest recall accuracy. These results do not support the peg hypothesis (Paivio, 1991; Paivio et al., 1994) and Perceptual Symbol Systems Theory (Barsalou, 1999), as it was expected that a greater number of sensory modality codes would enhance recall.

Analysis of Recognition Data

In the recognition task, the effect of the assessment group was not statistically significant: $F(2,70) = 2.75$, $p = .07$, $\eta^2 = .075$. Consistent with recall tasks, the most accurate recognition was recorded for the vividness group, while the other two groups achieved similar accuracy (Table 7). Contrary to recall tasks, we recorded the NoM effect: $F(2,67) = 9.92$, $p < .001$, $\eta^2 = .228$, where the abstract words had significantly lower hit rates (corrected for false alarms) compared to many-modalities (*Mdiff* = -.058, $p = .008$) and compared to few-modalities nouns (*Mdiff* = -.074, $p < .001$). Table 7 presents recognition accuracy across the NoM factor. There was no significant difference between the two groups of concrete nouns. No significant interaction was found between the assessment group and the NoM factor: $F(4,136) = .655$, $p = .624$, $\eta^2 = .019$. Figure 3 depicts the recognition accuracy of different word categories and assessment groups.

Table 6
Recognition accuracy (hits corrected for false alarms) across the assessment groups

	Mean	SE	N
Concreteness	.709	.034	25
Context availability	.719	.035	23
Vividness	.814	.035	23

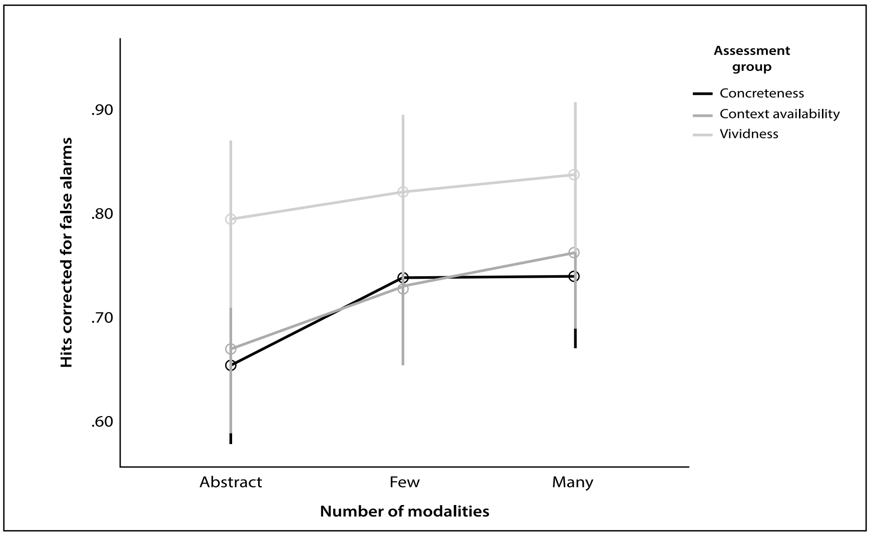
Note. SE = standard error.

Table 7
Recognition accuracy (hits corrected for false alarms) by the NoM word groups

	Mean	SE	N
Concreteness	.709	.034	25
Context availability	.719	.035	23
Vividness	.814	.035	23

Note. SE = standard error.

Figure 3
The interaction of assessment group and NoM in recognition task.



Note. The vertical lines denote 95% confidence intervals of the group means.

Discussion and Conclusion

Our study took a unique approach to explore the concreteness effect (e.g., Kounios & Holcomb, 1994; Paivio, 1991; Paivio et al., 1994; Popović Stijačić & Filipović Đurđević, 2015, 2022) in different memory tasks: free recall, cued recall, and recognition tasks. By varying the incidental encoding phase, we aimed to investigate the Depth of Processing Theory (Craik & Lockhart, 1972) and its influence on the retrieval of abstract and concrete words divided into those whose meaning can be experienced with few sensory modalities and those that can be experienced with many sensory modalities.

We partially confirmed the assumptions concerning the DoP Theory (Craik & Lockhart, 1972). As hypothesised, the participants spent most of the encoding time estimating the vividness of mental images of target words, which was followed by the highest recall and recognition accuracy. However, the encoding time could not explain the better reproduction and recognition for the concreteness assessment group compared to the context availability group since there was no difference in encoding time between these two groups. This result implies that the depth of processing cannot be solely operationalised via encoding time. Still, as the recall was consistently higher for the vividness group, we may conclude that assessing vividness requires additional mental effort compared to assessing concreteness and context availability since the participants were instructed to create mental images of a given word.

The most prominent finding was the opposite direction of the concreteness effect in the dependency of different memory tasks. In this study, we extended the traditional definition of concreteness and operationalised perceptual richness as the number of sensory modalities through which a concept can be experienced. We divided concrete words into two groups: the first group contained nouns that refer to concepts that can be experienced with few senses, and the other contained concepts experienced with many senses. Such manipulation enabled us to test the peg hypothesis (Paivio, 1991; Paivio et al., 1994), extended with the assumptions of the Perceptual Symbol Systems Theory (Barsalou, 1999, 2002). Our results contrast with the peg hypothesis, considering that both groups of concrete words were averaged for imageability and concreteness; thus, the groups differed only in perceptual richness, which did not enhance the reproduction of many-modalities words.

Some researchers (e.g., Anderson & Peterson, 2022; Parker & Dagnall, 2009) have hypothesised that Visual Working Memory (VWM) is engaged during the processing of concrete words due to the mental visualisation processes typically associated with these words. To test this assumption, they applied a Dynamic Visual Noise (DVN) paradigm while participants learned lists of concrete and abstract words during the encoding phase. DVN involves a continuously changing visual pattern that participants view while simultaneously listening to the words they are asked to remember (Anderson & Peterson, 2022; Parker & Dagnall, 2009).

In studies employing this paradigm, the authors reported a reversed concreteness effect, that is, inconsistent performance across memory tasks (e.g., free recall and recognition), attributed to the presence of DVN. They concluded that visual interference impaired participants' ability to utilise visual working memory (utiliseVWM), thereby disrupting the encoding of concrete words. In contrast, the encoding of abstract words, which rely less on visual imagery, remained largely unaffected by DVN (Anderson & Peterson, 2022; Parker & Dagnall, 2009).

The DVN paradigm can partially explain our results, as abstract words showed the highest recall accuracy but only in free recall. In cued recall, equal recall accuracy was achieved for the abstract and few-modalities words, compared to many-modalities nouns. On the other hand, in the recognition task, the hit rate (corrected for false alarms) was highest for the many-modalities word group. Therefore, the concreteness effect remained intact. One possible explanation is that concrete words are more distinct than abstract ones, leading to better recognition accuracy. This explanation is grounded in a recent study which found that concrete words exhibit properties of *typical word form*, making them distinct at the word form level (*concreteness form typicality*) and easier to process in various word processing tasks (Kearney et al., 2024). However, we did not manipulate this variable; thus, future studies should explore this explanation. Furthermore, as we did not include a control group (i.e., a group without the mental rotation task), we cannot unequivocally conclude that the mental rotation task prevented mental visualisation during the consolidation of memory traces.

The novelty of this study lies in the fine grain of the word concreteness. Namely, none of the previously mentioned studies varied the perceptual richness of words or investigated the possible influence of additional perceptual cues related to word meanings as additional signs during recall or recognition. Furthermore, the advantages of our research included incidental learning, which allowed us to control the strategic imagery during the encoding. Further, we loaded VWM during the memory consolidation phase, while, in previous studies, such manipulation was applied during the encoding or retrieval phase.

To summarise, our study demonstrated the reversed concreteness effect in incidental recall tasks under the load of Visual Working Memory during memory consolidation, as well as the concreteness effect in recognition tasks under the same conditions. By manipulating the perceptual richness of concrete nouns, we can reject the peg hypothesis, as additional perceptual cues do not enhance the retrieval of concepts that can be experienced through more than two senses. Future studies should include a control group in the design, match the word form typicality of abstract and concrete words, and examine different working memory loads.

Acknowledgements

We would like to express our sincere gratitude to the Faculty of Media and Communications for providing us with the necessary support and to the participants for taking part in this research.

References

Anderson, B. A., & Peterson, S. A. (2022). Concreteness and levels of processing: a test of the dual-coding hypothesis using dynamic visual noise. *Memory*, *30*(10), 1405–1420. https://doi.org/10.1080/09658211.2022.2120622

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Doest, L. T., & Semin, G. (2005). Retrieval contexts and the concreteness effect: Dissociations in memory for concrete and abstract words. *European Journal of Cognitive Psychology*, *17*(6), 859–881. https://doi.org/10.1080/09541440540000031

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 351-360. https://doi.org/10.1016/S0022-5371(66)80044-0

Hamilton, M., & Rajaram, S. (2001). The concreteness effect in implicit and explicit memory tests. *Journal of Memory and Language, 44*(1), 96–117. https://doi.org/10.1006/jmla.2000.2749

IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

Kearney, E., McMahon, K. L., Guenther, F., Arciuli, J., & De Zubicaray, G. I. (2024). Revisiting the concreteness effect: Non-arbitrary mappings between form and concreteness of English words influence lexical processing. *Cognition*, *254*, 105972. https://doi.org/10.1016/j.cognition.2024.105972

Kostić, Đ. (1999). Frekvencijski rečnik savremenog srpskog jezika. Tom I – VII. Institut za eksperimentalnu fonetiku I patologiju govora, Beograd i Laboratorija za eksperimentalnu psihologiju Filozofskog fakulteta u Beogradu.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 804. https://doi.org/10.1037/0278-7393.20.4.804

Leiner, D. J. (2023). SoSci Survey [Computer software]. https://www.soscisurvey.de

Marks D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482–488. https://doi.org/10.1037/h0045106

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*(3), 255. https://doi.org/10.1037/h0084295

Paivio, A., Walsh, M., & Bons, T.A. (1994). Concreteness effects on memory: when and why? *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 1196–1204. https://psycnet.apa.org/doi/10.1037/0278-7393.20.5.1196

Parker, A., & Dagnall, N. (2009). Concreteness effects revisited: the influence of dynamic visual noise on memory for concrete and abstract words. *Memory*, *17*(4), 397–410. https://doi.org/10.1080/09658210902802967

Popović-Stijačić, M., & Filipović Đurđević, D. (2015). Number of sensory modalities through which a concept can be experienced: effect on recall. *Primenjena psihologija*, *8*(3), 335-352. https://doi.org/10.19090/pp.2015.3.335-352

Popović Stijačić, M. (2021). The relevance of perceptual information in word processing – the perspective of embodied cognition [University of Novi Sad]. University of Novi Sad. https://www.cris.uns.ac.rs/DownloadFileServlet/Disertacija161762163851549.pdf?controlNumber=(BISIS)117683&fileName=161762163851549.pdf&id=17699&licenseAccepted=true

Popović Stijačić, M., & Filipović Đurđević, D. (2022). Perceptual richness and its role in free and cued recall. *Primenjena psihologija*, *15*, 255-381. http://dx.doi.org/10.19090/pp.v15i3.2400

Romani, C., Mcalpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, *61*(2), 292–323. https://doi.org/10.1080/17470210601147747

Schwanenflugel, P. J. (1991). Contextual constraint and lexical processing. *Advances in psychology*, *77*, 23-45. https://doi.org/10.1016/S0166-4115(08)61528-9

Schwanenflugel, P.J., Akin, C. & Luh, WM (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition, 20*, 96–104. https://doi.org/10.3758/BF03208259

Schwanenflugel, P. J. & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, *24*(1), 114-126.https://psycnet.apa.org/doi/10.2307/748013

Vandendaele, A., & Grainger, J. (2022). Now you see it, now you don't: Flanker presence induces the word concreteness effect. *Cognition*, *218*, 104945. https://doi.org/10.1016/j.cognition.2021.104945

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*(6), 1024–1037. https://doi.org/10.1162/08989290051137558

APPENDIX

Table A
The list of target and cue stimuli

Serbian		English translation	
Target	**Cue**	**Target**	**Cue**
DOSADA	avantura	boredom	adventure
PEDALA	bicikl	pedal	bicycle
SLIČNOST	bliskost	similarity	closeness
KLJUČ	brava	key	lock
PEPELJARA	dim	ashtray	smoke
LIPA	drvo	linden (tree)	tree
MRAVINJAK	gužva	anthill	crowd/jam
MUVA	insekt	fly	insect
LAŽ	istina	lie	truth
DIVOTA	krasota	delight	beauty/splendour
DAVANJE	krv	giving	blood
ANĐEO	lepota	angel	beauty
KARMIN	lice	lipstick	face
VOĐA	lider	leader	leader
POKLOPAC	lonac	lid	pot
MODRICA	masnica	bruise	bruise
POŠTENJE	moral	honesty	morality
OSTRVO	more	island	sea
LOPATA	motika	shovel	hoe (garden tool)
GENIJE	naučnik	genius	scientist
KOŠNICA	pčela	beehive	bee
MEDUZA	pipak	jellyfish	tentacle
BUNDEVA	pita	pumpkin	pie
PESAK	plaža	sand	beach
JELKA	poklon	fir tree / Christmas tree	gift
PONUDA	potražnja	offer	demand
PAUČINA	prašina	cobweb	dust
OBIČAJ	praznik	custom	holiday
VILICA	proteza	jaw/fork (depending on context)	denture/prosthesis
VERNIK	religija	believer	religion
LUSTER	sijalica	chandelier	light bulb
SLAMČICA	sok	straw (drinking)	juice
BANDERA	stub	pole	pillar/pole
SVETLOST	sunce	light	sun
SUMRAK	suton	twilight	dusk/twilight
TEG	vaga	weight	scale/balance
PLJUVAČKA	varenje	saliva	digestion

PROLEĆE	vesnik	spring	herald/messenger
MAHOVINA	vlaga	moss	moisture/humidity
ARMIJA	vojska	army	army
SVEĆA	vosak	candle	wax
ZVONCE	vrata	bell	door(s)
DOBROTA	vrlina	kindness	virtue
ROLETNA	zavesa	blind / roller blind	curtain
KOPNO	zemlja	land	earth/land/soil