Multiple object categorization and effect of spatial frequencies

Alvydas Soliunas Corresp., [1] , **Aleksandras Pleskaciauskas** [1] , **Algis Daktariunas** [1]

[1] Institute of Biosciences, Vilnius University, Vilnius, Lithuania

Corresponding Author: Alvydas Soliunas
Email address: alvydas.soliunas@gf.vu.lt

A process of interaction between objects and scene is widely investigated but much less attention is paid to the interaction between objects in multiple objects stimuli. In psychophysical experiment, we presented one, two, or three visual objects simultaneously for 100 ms and then asked subjects to answer whether objects belong to the same category (Experiments 1 and 2), or whether afterwards presented probe-word signify an object that was presented (Experiments 3 and 4). Interestingly, performance accuracy and reaction time did not depend on the number of objects if they belonged to the same category, but performance deteriorated when more categories were presented. Filtering out high or low spatial frequencies did not affect performance peculiarities of the objects of the same or different categories. The findings support assumption that visual objects of the same category could be identified simultaneously but the different categories are identified successively.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26666v1 | CC BY 4.0 Open Access | rec: 13 Mar 2018, publ: 13 Mar 2018

Multiple object categorization and effect of spatial frequencies

Alvydas Soliunas[1], Aleksandras Pleskaciauskas[1], Algis Daktariunas[1]

[1] Instuitute of Biosciences, Vilnius University, Vilnius, Lithuania

Corresponding Author: Alvydas Soliunas

Saulėtekio al. 7, LT-10257 Vilnius, Lithuania

Email address: alvydas.soliunas@gf.vu.lt

Abstract

A process of interaction between objects and scene is widely investigated but much less attention is paid to the interaction between objects in multiple objects stimuli. In psychophysical experiment, we presented one, two, or three visual objects simultaneously for 100 ms and then asked subjects to answer whether objects belong to the same category (Experiments 1 and 2), or whether afterwards presented probe-word signify an object that was presented (Experiments 3 and 4). Interestingly, performance accuracy and reaction time did not depend on the number of objects if they belonged to the same category, but performance deteriorated when more categories were presented. Filtering out high or low spatial frequencies did not affect performance peculiarities of the objects of the same or different categories. The findings support assumption that visual objects of the same category could be identified simultaneously but the different categories are identified successively.

Introduction

We are living in an environment where many objects are seen at the same time. When we are in unfamiliar environment (town or building) and glance at some direction, it seems for us that we instantly perceive all environment and all objects there. But this is not true. Conscious perception of objects is successive process, about 3-5 items per second (Koch, 2004; Del Cul et al., 2007; Sergent et al., 2005; Madl et al., 2011). In a more controlled situation when objects are presented sequentially at fixation, the speed of object identification could reach 8 items per second (Potter 1976). Furthermore, objects normally are not recognized in isolation.

One kind of interaction is influence of scene as a context on recognition of objects in scene. There are many studies that demonstrate facilitating context effect on object recognition in scenes (Biederman et al., 1982; Davenport & Potter, 2004; Joubert et al., 2007). Scene identification influences recognition of objects in the scene, and on the contrary, recognition of objects influences identification of scene (Joubert et al., 2007, 2008; Mack & Palmeri, 2010).

Another kind of interaction is object-to-object interaction. Our work addresses namely this kind of interaction. We are interested in interaction between objects in the process of categorization of visual objects when several objects are presented simultaneously. To simplify the situation, we chose brief 100 s presentation conditions, where all objects should be analyzed without saccadic eye movement. If the objects are small enough and they fit in an area of central vision, we can expect two scenarios in respect of simultaneous versus successive analysis of objects: (i) objects are recognized or categorized successively; (ii) objects are recognized or categorized in parallel; (iii) hybrid scenario, when objects basically are analyzed in parallel, but some are recognized faster than others and therefore they could influence categorization of other objects. Simultaneous categorization supposes independent categorization of objects, as opposite to successive categorization that permit dependent categorization when recognition of the particular object influences the recognition of other objects.

Some studies demonstrate pre-attentive categorization of objects in a scene (Li et al., 2002; Evans & Treisman, 2005, Poncet et al., 2012; Rousselet et al., 2002) that could suggest the parallel categorization of several objects. Rousselet at al. (2002) demonstrate that even complex scenes presented simultaneously can be processed in parallel. Other authors state that the scenario could depend on the level of categories. Some studies show that categorization of objects on basic level require more time than categorization on superordinate level (Mace et al., 2009). Gronau et al. (2008; see also Aucland et al., 2007) demonstrate importance of semantic and spatial relations between objects in a study where two semantically related or unrelated objects were presented in congruent or incongruent spatial relation.

Back to our research, we can predict following results of categorization of simultaneously presented visual objects: in case of scenario (i), there should be direct dependence of reaction time on the number of objects; in case of scenario (ii), there should be different reaction time of the categorization of particular objects, depending on the categories or similarity of surrounding objects.

Experiment 1

The goal of the first experiment was to find whether object categorization depends on diversity of objects in the stimulus, i.e. whether there is difference when objects belong to the same category or to the different categories.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26666v1 | CC BY 4.0 Open Access | rec: 13 Mar 2018, publ: 13 Mar 2018

Material and Methods

Subjects

Forty three volunteer students from Vilnius University (eleven males and 32 females, 20-23 years of age) took part in this experiment. Each subject had normal or corrected-to-normal vision and had no prior experience with psychophysical testing of similar nature. They were naive to the goals of the experiment and signed an informed consent approved by Vilnius Region Ethics Commitee of Biomedical Research (approval No.158200-13-578-173). All subjects took part in one experimental session.

Stimuli

We used grayscale version of computer-generated images from TarrLab Object Databank (stimulus images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, http://www.tarrlab.org/). Eleven categories of objects were presented (bottle, brush, chair, sofa, desk, table, clock, cup, telephone, vase, pot), and there were four objects of each category, i.e. altogether 44 objects.

Stimulus consisted of fixation point and two or three objects located around it. Three objects were evenly distributed in a square area of 8° x 8° around fixation point. When two objects were presented, they were distributed in such a manner as if a third object would be present. There were five types of stimuli (Fig.1): 1) “1-2” stimuli (two different objects of one category); 2) “1-3” stimuli (three different objects of one category); 3) “2-2” stimuli (two objects of two categories); 4) “2-3” stimuli (three objects of two categories); 5) “3-3” stimuli (three objects of three categories). Altogether 330 stimuli (66 of each type) were presented in a random order. As there were 44 different objects, each object occurred approximately in 20 stimuli. There were 6-8 different 3D orientations for each object. Different orientations were randomized among stimuli. White 8° x 8° stimulus area had no borderlines and was presented on white screen background.

<insert Fig.1 about here>

Procedure

Experiment was performed in a room with natural daylight illumination. Stimulus presentation and data registration were under control of a computer equipped with 19-inch CRT monitor (1024 x 768 resolution and 85 Hz frame-rate), standard keyboard and a *Stimscope (© R. Zoontjens 1997)* experiment generator running under Windows OS.

Subject's head was not fixed but they were instructed to hold the same distance, about 65 cm, from display during experiment.

Examples of all eleven categories of objects and the names of categories were presented to subjects for a few minutes before experiment. Then subjects performed practice block of 15 trials. An event sequence within a trial is shown in Fig. 2. Fixation point was presented at the center of screen for 300 ms and the subjects were asked to keep their eyes focused on fixation point during the test stimulus presentation. Fixation point was followed by a 100 ms blank interval and then a test stimulus was displayed for 100 ms. We used backward masking procedure, i.e. test stimulus was masked by a 8° x 8° square of chaotic pattern (see Fig .2) for 300 ms. Subjects were instructed to press the [V] key on a keyboard if they guessed that all objects in the stimulus belonged to the same category, or to press the [N] key, if they belonged to different categories. Subjects had four seconds to make their decision. The response initialized the next trial with 100 ms delay. The whole experiment session lasted about 40 minutes.

<insert Fig.2 about here>

As there were large individual differences, the response accuracy and RT data were normalized for statistical analysis. Accuracy data were normalized in respect to 88 %, and the RT data were normalized in respect to 600 ms. These values were close to grand mean for all subjects. Normalization formula was $T_{ni} = T_i + (600 – T_m)$ for RT data and $P_{ni} = P_i + (88 – P_m)$ for accuracy data. T_{ni} and P_{ni} were normalized values, T_i and P_i – raw values, T_m and P_m – means for given subject.

Results.

Results of Experiment 1 are presented in Figure 3 under the column "Original". The data are presented in a diagram of linear type because of easier visual interpretation. Because of incomplete block design, the two ANOVA were performed: (i) three-way ANOVA for factors of sameness (same vs different categories), categories (one, two, or three categories) and gender; (ii) two-way ANOVA for factors of categories (one, two, or three categories) and objects (two or three objects). The only significant factor was the categories: $F(2,209) = 16.01$, $p < 0.0001$ for RT and $F(2,209) = 7.35$, $p < 0.001$ for response accuracy. According to Duncan post hoc test, RT did not differ whether objects belonged to one or to two categories (respectively 604 ms 643 ms, $p = 0.262$), but the RT was shorter (564 ms, $p < 0.0001$) when objects of three categories were presented. Accuracy was significantly different ($p < 0.001$ in all cases) for all three cases: it was highest for the three-category stimuli (92.1 %), middle for one-category stimuli (88.9 %) and lowest for two-category stimuli (85.0 %).

<insert Fig.3 about here>

There was significant interaction between two factors - categories and objects: $F(1,210) = 20.36$, $p < 0.0001$ for RT and $F(1,210) = 23.51$, $p < 0.0001$ for accuracy. This interaction is clearly visible in fig. 3. We can make several conclusions on this:

1) For one-category stimuli, there was no difference in RT whether two or three objects were presented ("1-2" vs "1-3" stimuli, 605 ms vs 606 ms, $p = 0.937$), but accuracy was slightly higher for three-objects stimuli (87.5 % vs 90.3 %, $p = 0.031$). This result supports the hypotheses of simultaneous categorization of objects of the same category.

2) When two objects were presented, there was no difference in accuracy magnitude whether they were of the same or different categories ("1-2" vs "2-2" stimuli, 87.5 % and 87.9 %, $p = 0.755$), but RT was slightly faster for the objects of different categories (605 ms vs 584 ms, $p = 0.036$). This result more agrees with successive categorization mode of objects, but simultaneous categorization mode cannot be rejected as there is possibility to correctly answer "Different" without identifying the second object: e.g. first object is identified as "a chair" and the second one as "not a chair" which is less demanding task than identifying exact category of second object.

3) When three objects were presented, there was no difference in accuracy magnitude whether they were of the same or different categories ("1-3" vs "3-3" stimuli, 90.3 %, vs 92.1 %, $p = 0.156$), but RT was faster for the objects of different categories (606 ms vs 564 ms, p <

0.0001). As there was no need to identify all three objects in case of "3-3" stimuli, the result more agrees with successive categorization mode of objects.

The "2-3" stimuli was special case: because they consisted of objects of the same category as well as objects of different categories, the task "same-different" was most difficult in this case and naturally the RT and accuracy were the worst.

Experiment 2

Results of Experiment 1 let us suppose the possibility of simultaneous categorization of visual objects that belong to the same category. It is possible that simultaneous categorization could be based primarily on low spatial frequencies. There are many works in scene perception that demonstrate that scene perception could be based on low spatial frequencies (or by magnocellular pathway) (Oliva & Schyns, 1997; Delorme et al., 2000, 2010; Oliva & Torralba, 2006; Bar, 2004; Thorpe, 2011). In Experiment 2, we tried to test the role of high and low spatial frequencies in categorization of multiple objects.

Experiment 2 was the replica of Experiment 1 with high-pass filtered (i.e. low spatial frequencies are eliminated) and low-pass filtered stimuli instead of original ones.

Material and Methods

Subjects

Twenty volunteer students from Vilnius University (ten males and ten females, 20-28 years of age) took part in one experimental session of this study. Each subject had normal or corrected-to-normal vision and had no prior experience with psychophysical testing of similar nature. They were naive to the goals of the experiment.

Stimuli

The same 330 stimuli from experiment 1 were used in this experiment. They were modified by filtering low or high spatial frequencies. The filters were applied from *ImageJ (Image Processing and Analysis in Java 2013)* image editor. High spatial frequencies were filtered (i.e. applied low-pass filter) with *Gaussian Blur* filter with sigma = 2 (Fig. 4). Lower spatial frequencies were filtered with *Find Edges* filter.

Three hundred thirty low-pass stimuli and 330 high-pass stimuli were presented separately and the presentation order of two blocks was counterbalanced among subjects.

All other conditions of stimuli presentation and procedures were the same as in Experiment 1.

<insert Fig.4 about here>

Results

Results of Experiment 2 are presented separately for low-pass (or Low SF) and high-pass (or High SF) stimuli in Figure 3. Significant main effects of ANOVA are presented in Table 1.

<insert Table.1 about here>

The number of categories was a significant factor in all four cases. On the contrary, the factor of the number of objects was significant only for RT of low-pass stimuli.

For low-pass stimuli, interaction between category and object factors was significant for accuracy data ($F(1,95) = 10.46$ $p < 0.01$), but was not significant for RT data (($F(1,95) = 1.489$ $p = 0.225$). We had similar situation for high-pass stimuli: $F(1,95) = 8.911$ $p < 0.01$ for accuracy and $F(1,95) = 2.915$, $p = 0.091$ for RT. It is possible that not significant interaction between two factors was due to high individual variability of RT data.

As in Experiment 1, we should stress some important results:

1) For one-category stimuli, there was no difference in RT and response accuracy for low-pass stimuli whether two or three objects were presented ("1-2" vs "1-3" stimuli, 582 ms vs 609 ms, p = 0.057 for RT, and 85.1 % vs 88.1 %, p = 0.141 for accuracy), and the same is true for high-pass stimuli ("1-2" vs "1-3" stimuli, 598 ms vs 608 ms, p = 0.433 for RT, and 84.7 % vs 87.3 %, p = 0.280 for accuracy). Altogether these findings basically correspond to the findings of Experiment 1 and support hypotheses of simultaneous categorization of objects of the same category.

2) When two objects were presented, there was no difference in RT whether they were of the same or different categories ("1-2" vs "2-2" stimuli, 582 ms vs 589 ms, p = 0,588 for low-pass stimuli, and 598 ms vs 590, p = 0.545 for high-pass stimuli) but accuracy was higher for the

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26666v1 | CC BY 4.0 Open Access | rec: 13 Mar 2018, publ: 13 Mar 2018

objects of different categories (85.1 % vs 89.7 %, $p = 0,030$ for low-pass stimuli, and 84.7 % vs 90.7 % $p = 0,017$ for high-pass stimuli). This corresponds to the findings of Experiment 1: "2-2" stimuli were classified more effectively as "Different" than "1-2" stimuli as "Same".

3) When three objects were presented, the RT was faster and accuracy was higher for the objects of different categories ("1-3" vs "3-3" low-pass stimuli: 609 ms vs 580 ms, $p = 0.045$ for RT, and 88.1 % vs 93.6 %, $p = 0.011$ for accuracy; "1-3" vs "3-3" high-pass stimuli: 608 ms vs 571 ms, $p < 0.01$ for RT, and 87.3 % vs 94.1 %, $p < 0.01$ for accuracy). And again, this corresponds to the findings of Experiment 1: "3-3" stimuli were classified more effectively as "Different" than "1-3" stimuli as "Same".

In summary, the results of Experiment 2 repeated the results of Experiment 1 and essentially there was no difference between performance on low-pass and high-pass stimuli.

Experiment 3

Findings of the first two experiments could suggest simultaneous identification of objects of the same category but the task for subjects, same-different task, did not required identification of all objects under all five conditions. In case of three objects of two or three categories, it was possible to answer "Different" after identification of only two objects if they were from different categories. To further investigate categorization of multiple objects we chose simple identification task: stimulus with multiple objects is followed by a probe-word and subject should answer whether a probed object was present in the stimulus. As the subject did not know what category word would be presented, he should identify all objects in the stimulus.

Material and Methods

Subjects

Six volunteer students from Vilnius University (three males and three females, 20-22 years of age) took part in one experimental session of this study. Each subject had normal or corrected-to-normal vision and had no prior experience with psychophysical testing of similar nature. They were naive to the goals of the experiment.

Stimuli

Stimuli were created from the same set of objects as in Experiment 1. There were six types of stimuli: 1) “1-1” stimuli (one object); 2) “1-2i” stimuli (two identical objects of one category); 3) “1-2d” stimuli (two different objects of one category); 4) “2-2” stimuli (two objects of two categories); 5) “2-3” stimuli (two different objects of one category and third object of other category); 6) “3-3” stimuli (three object of three categories). An object in “1-1” stimulus was located not in the center of 8° x 8° stimulus area but near fixation point at random location. One hundred thirty two stimuli (22 of each type) were repeated four times, thus 528 stimuli were presented in random order in one experimental session.

Procedure

Experiment was controlled by *E-Prime* v.2.0 (© *Psychology Software Tools Inc*.) experiment generator running under Windows OS. Practice block consisted of 24 trials, i.e. four trials of each stimulus type. An event sequence within a trial was similar to the procedure in Experiment 1 (Fig. 2), and only difference was a probe-word which was presented after masking pattern. The probe-word was a name of a category written in lowercase Ariel font, 2° height. Subjects had to decide whether an object defined by a probe-word was presented or not on a given trial by pressing the [1] or [2] key on the right side of a keyboard. One half of subjects received the instruction to press the [1] for “Yes” answer and [2] for “No” answer, whereas another half received inverse instruction. Other conditions of the experiment were the same as in Experiment 1.

Results.

Results of Experiment 3 are presented in Figure 5 on the left side. There was no statistical difference between performance results for the “1-2i” and “1-2d” stimuli and we merged these results into one group “1-2” for further analysis. It was not possible to conduct one ANOVA for all factors because of incomplete design of experiment, therefore we conducted separate ANOVA’s for the factors of presence (object presented or not presented), number of objects (one, two or three), and number of categories (one, two or three). There was significant main effect of the factor of presence for RT data (($F(1,10) = 14.14$, $p < 0.01$) indicating the faster response when probed object was presented (645 ms vs 755 ms). Main effect of objects ($F(2,33) = 28.63$, $p < 0.0001$ for RT and $F(2,33) = 15.49$, $p < 0.0001$ for accuracy) indicate that the higher the number of objects, the slower RT and the lover accuracy was (601 ms vs 654 ms vs 817 ms and 94.4 % vs 92.6 % vs 83.9 %, respectively for the stimuli of one, two, and three objects). Main effect of categories ($F(2,33) = 190.6$, $p < 0.0001$ for RT and $F(2,33) = 48.44$, $p < 0.0001$ for accuracy) indicate that the higher the number of categories, the slower RT and the lover accuracy was (598 ms vs 785 ms vs 834 ms and 94.9 % vs 87.5 % vs 80.4 %, respectively for the stimuli of one, two, and three categories).

<insert Fig.5 about here>

It is obvious that the dependence of performance effectiveness is stronger and is more clearly expressed on the number of categories than on the number of objects (see Fig 5). The statement “the more categories, the poorer performance” is correct irrespectively of the number of objects, except one case - insignificant difference between RT on “2-3” and “3-3” stimuli. On the contrary, the statement “the more objects, the poorer performance” basically is not correct, because it depends on the number of categories, i.e. we have interaction between the factor of categories and the factor of objects. This interaction could be demonstrated by the following results:

1) For one-category stimuli, there was no difference in RT and accuracy whether one or two objects were presented (“1-1” vs “1-2” stimuli, 601 ms vs 587 ms, and 94.4 % vs 95.2 %).

2) For two-category stimuli, there was no difference in RT and accuracy whether two or three objects were presented (“2-2” vs “2-3” stimuli, 770 ms vs 800 ms, and 87.6 % vs 87.4 %).

3) For two-object stimuli, RT was faster and accuracy higher when the objects belonged to the same category (“1-2” vs “2-2” stimuli, $p < 0.0001$ for RT and $p < 0.001$ for accuracy on Duncan post hoc test).

4) For three-object stimuli, there was no statistical difference in RT (“2-3” vs “3-3” stimuli, 801 ms vs 804 ms), but accuracy was lower for three-category stimuli (87.4 % vs. 80.4 %, $p < 0.001$).

Taken together these findings support the hypotheses of simultaneous categorization of objects of the same category.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26666v1 | CC BY 4.0 Open Access | rec: 13 Mar 2018, publ: 13 Mar 2018

Experiment 4

Experiment 4 was replica of Experiment 3 with high-pass filtered and low-pass filtered stimuli instead of original ones. Filtering high and low spatial frequencies did not affect same-different task performance in Experiment 2 but the task in Experiment 3 was different therefore the role of high and low spatial frequencies in task performance could differ.

The low-pass stimuli (132 stimuli repeated twice) and high-pass stimuli were presented separately and presentation order of two blocks was counterbalanced between subjects. Thirty-one subjects (25 males and six females, 20-23 years of age) took part in one experimental session. All other conditions of stimuli presentation and procedures were the same as in Experiment 3.

Results of Experiment 4 are presented in Figure 3. There were significant effects of: the number of objects (low-pass stimuli: $F(2,183) = 24.67$, $p < 0.0001$ for RT and $F(2,183) = 18.26$, $p < 0.0001$ for accuracy; high-pass stimuli: $F(2,183) = 24.33$, $p < 0.0001$ for RT and $F(2,183) = 15.61$, $p < 0.0001$ for accuracy); the number of categories (low-pass stimuli: $F(2,183) = 189.7$, $p < 0.0001$ for RT and $F(2,183) = 338.1$, $p < 0.0001$ for accuracy; high-pass stimuli: $F(2,183) = 224.3$, $p < 0.0001$ for RT and $F(2,183) = 339.2$, $p < 0.0001$ for accuracy). Interaction of these two factors was not significant. We emphasized four findings in Experiment 3 that characterize how multiple objects categorization depends on the number of categories and the number of objects. As these four findings were repeated in Experiment 4 (see Fig.5), there is no need to describe them again. There was only one small difference which even strengthens the findings of all experiment: “2-3” stimuli were processed faster than “3-3” stimuli. No differences were observed between performance on low-pass and high-pass stimuli (spatial frequency factor: $F(1,370) = 0.89$, $p = 0.346$ for RT and $F(2,183) = 1.83$, $p = 0.177$ for accuracy).

General discussion

The first question we address in our study, is whether processing of several objects presented simultaneously is running independently or identification of particular object depends on identification of other objects. Another aspect of this question is – whether multiple isolated objects are categorized successively or in parallel.

In studies of scene perception, we can see plenty of evidence that two processes, the scene recognition (gist) and objects recognition in the scene, are running interdependently. In our study, we had only isolated objects without scene context. Results of all four experiments (each with four different group of subjects) primarily support an idea of parallel processing of separate objects presented simultaneously, but this concerns only objects of the same category. This statement is based on the following findings: 1) there was no difference between task performance effectiveness of one and two objects if they belong to the same category (Experiments 3 and 4); 2) there was no difference between task performance effectiveness of two and three objects if they belong to the same category (Experiments 1 and 2); 3) there was no difference between task performance effectiveness of two and three objects if they belong to two categories (Experiments 3 and 4). It looks like this: when we recognize one chair, all other chairs in the visual field are recognized instantly. Moreover, it does not depend on physical similarity of objects (objects of the same category differ by global shape, texture, local features, orientation). In Experiment 2 and 4 there were two types of two-objects-one-category stimuli: two identical objects ("1-2i" stimuli) and two different objects of the same category ("1-2d" stimuli). We did not find any difference in performance for these two types of stimuli.

On the other hand, results of our experiments support a statement of successive categorization process of simultaneously presented isolated objects if they belong to different categories. Response time and accuracy directly depended on the number of categories in Experiment 3 and 4: the more categories were presented, the lower performance accuracy and the longer response time was. Basically such results correspond to the scenario (iii) that was mentioned in Introduction, i.e. hybrid scenario.

Our results raise a question - what makes difference between within-category features and between-category features. We tested the hypothesis that parallel recognition of objects of the same category could be based on low spatial frequencies and we did not find any essential differences in task performance for low-pass (blurred) stimuli and high-pass (contour) stimuli. This could suggest that both low and high spatial frequencies are involved in categorization of visual objects. On the other hand, we cannot deny a possibility that applied filters of spatial frequencies in our experiments were not powerful enough. In Collin and McMullen (2005) study low-pass filtering impaired categorization on subordinate level but had little effect on basic level category verification. It should be noted that our finding primarily concerns categories of basic level and our experiments were not designed to investigate categorization modes (parallel vs successive) of superordinate or subordinate levels.

Could we conclude that the visual objects of the same category are identified simultaneously in multiple object situation? Our findings (firstly three ones mentioned above) just support such a statement but do not prove it, because there could be alternative interpretations. One of them relates to the number of choices under different conditions of experiment. If there are two objects of different categories and subject identify both categories, he needs to remember two categories and after probe-word presentation he should search probed category between two memory items. If there are three objects of different categories and subject identify all categories, he should search probed category between three memory items. The bigger memory set size, the longer response time is. This effect was demonstrated in numerous studies in different situations. Such effect could take place in Experiments 3 and 4, but not in Experiments 1 and 2, because the short-term memory was unnecessary in the task of these experiments.

We can hypothesize about existence of a special mechanism of visual system which enables to recognize objects of “active“ category in parallel mode. Let us imagine situation when we see an audience of 100 people. You can get two tasks in this situation: (i) to find your colleague N., i.e. particular person; (ii) to find your colleague, i.e. there is one of your twenty colleagues of the work, but you do not know who. It is obvious that the first task is much easier and you will do it more quickly. In this case active category is a representation of particular person. The same is with well-known priming effect when top-down processes activate detection or identification of particular features of primed (i.e. activated) object. In experiments presented in this study, active category could be the category that is identified firstly between multiple objects. Identified category could enable simultaneous identification of all other objects of this category. Later we identify another category and this new category become active category. Our presented investigation does not allow to strictly validate neither existence of such mechanism nor directly matching of our findings with priming effect in other investigations.

Conclusions

In psychophysical experiments we investigated identification of visual objects presented simultaneously. All results of experiments were analyzed seeking to distinguish two visual processing modes – simultaneous versus successive identification of objects. Findings of the same-different task were contradictory. Some of them supported simultaneous identification mode, others supported successive identification mode. Results of the second task, where subjects had to detect probed category in stimulus with multiple objects, were more consistent with simultaneous identification of objects if they belonged to the same category. In case the objects belonged to different categories, they were identified successively.

References

Auckland ME, Cave KR, Donnelly N. 2007. Nontarget objects can influence perceptual processes during object recognition. *Psychonomic Bulletin and Review* 14(2):332-337.

Bar M**.** 2004. Visual objects in context. *Nature Reviews Neuroscience* 5(8):617-629. DOI: 10.1038/nrn1476.

Biederman I, Mezzanote RJ, & Rabinowitz JC. 1982. Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology* 14:143-177.

Collin CA, McMullen PA. 2005. Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization. *Perception & Psychophysics* 67(2):354-364.

Davenport JL**,** Potter MC**.** 2005. Scene Consistency in Object and Background Perception. *Psychological Science* 15:559-564. DOI: 10.1111/j.0956-7976.2004.00719.x.

Del Cul A, Baillet S, Dehaene S. 2007. Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology* 5(10):e260. DOI: 10.1371/journal.pbio.0050260.

Delorme A, Richard G, Fabre-Thorpe M. 2000. Ultra-rapid categorisation of natural images does not rely on colour: a study in monkeys and humans. *Vision Research* 40:2187–2200.

Delorme A, Richard G, Fabre-Thorpe M. 2010. Key visual features for rapid categorization of animals in natural scenes. *Frontiers in Psychology* 1:1–13. DOI: 10.3389/fpsyg.2010.00021.

Evans KK, Treisman A. 2005. Perception of objects in natural scenes: Is it really attention free. *Journal of Experimental Psychology: Human Perception and Performance* 31(6):1476–1492. DOI: 10.1037/0096-1523.31.6.1476.

Fabre-Thorpe M. 2011. The Characteristics and Limits of Rapid Visual Categorization. *Frontiers in Psychology* 2:243. DOI: 10.3389/fpsyg.2011.00243.

Gronau N, Neta M, Bar M. 2008. Integrated Contextual Representation for Objects' Identities and Their Locations. *Journal of Cognitive Neuroscience* 20(3):371–388. DOI: 10.1162/jocn.2008.20027.

Joubert OR, Rousselet GA, Fize D, Fabre-Thorpe M. 2007. Processing scene context: fast categorization and object interference. *Vision Research* 47(26):3286-97. DOI: 10.1016/j.visres.2007.09.013.

Joubert OR, Fize D, Rousselet GA, Fabre-Thorpe M. 2008. Early interference of context congruence on object processing in rapid visual categorization of natural scenes. *Journal of Vision* 8(13):11.1-18. DOI: 10.1167/8.13.11.

Koch Ch. 2004. *The Quest for Consciousness: A Neurobiological Approach*. Roberts & Co.

Li F, VanRullen R, Koch C, Perona P. 2002. Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America* 99(14):9596–9601. DOI: 10.1073/pnas.092277599.

Mace M J, Joubert O R, Nespoulous J L, Fabre-Thorpe M. 2009. The time-course of visual categorizations: You spot the animal faster than the bird. *PLoS One* 4:e5927. DOI: 10.1371/journal.pone.0005927.

Mack ML, Palmeri TJ. 2010. Modeling categorization of scenes containing consistent versus inconsistent objects. *Journal of Vision* 10(3):11.1-11. DOI: 10.1167/10.3.11.

Madl T, Baars BJ, Franklin S. 2011. The Timing of the Cognitive Cycle. *PLoS ONE* 6(4):e14803. DOI: 10.1371/journal.pone.0014803.

Oliva A, Schyns PG. 1997. Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology* 34(1):72-107. DOI: 10.1006/cogp.1997.0667.

Oliva A, Torralba A. 2006. Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research* 155(2):23-36. DOI: 10.1016/S0079-6123(06)55002-2.

Poncet M, Reddy L, Fabre-Thorpe M. 2012. A need for more information uptake but not focused attention to access basic-level representations. *Journal of Vision* 12(1):15. DOI: 10.1167/12.1.15.

Potter MC. 1976. Short term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory* 2:509-522.

Rousselet GA, Fabre-Thorpe M, Thorpe SJ. 2002. Parallel processing in high-level categorization of natural images. *Nature Neuroscience* 5:629–630. DOI: 10.1038/nn866.

Sergent C, Baillet S, Dehaene S. 2005. Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience* 8(10):1391-400. DOI: 10.1038/nn1549.

Figure 1

Examples of stimuli used in Experiment 1 and Experiment 2.

Top row (“Original”) represents Experiment 1, middle and bottom rows – Experiment 2 (Low SF – low-pass stimuli, High SF – high-pass stimuli). Columns represent five types of stimuli: 1-2 – one category, two objects; 2-2 – two categories, two objects; 1-3 – one category, three objects; 2-3 – two categories, three objects; 3-3 – three categories, three objects.

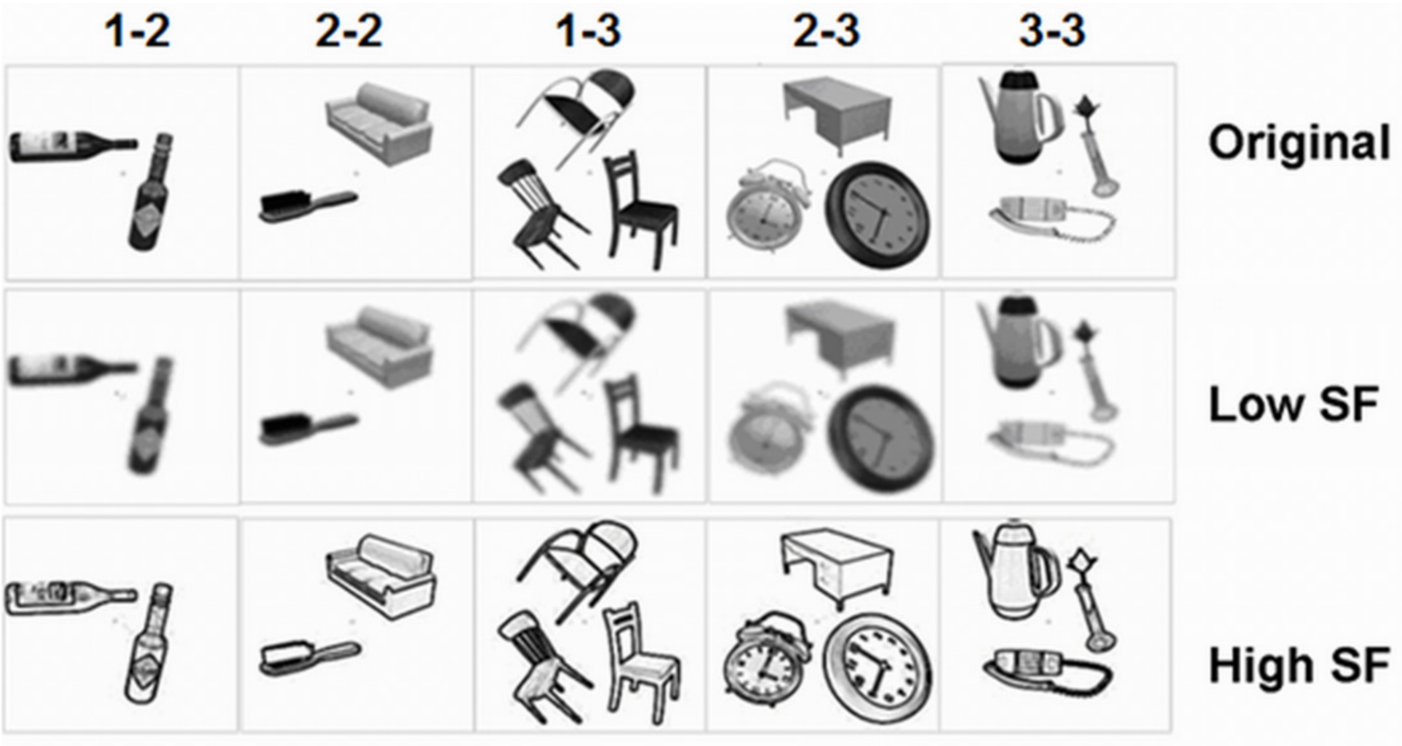
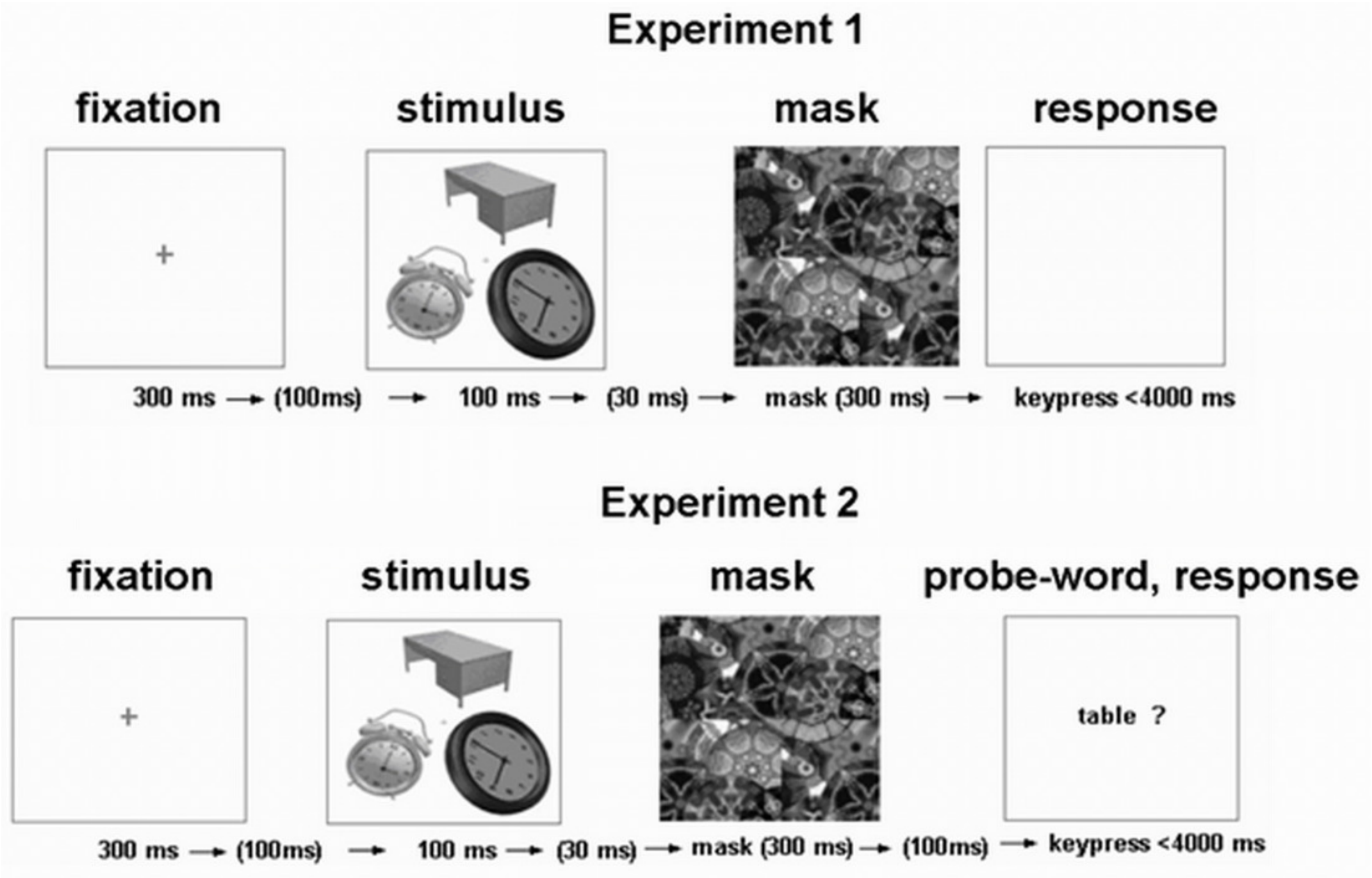


Figure 2

Stimuli presentation sequences within a trial in Experiments 1 and Experiment 3.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26666v1 | CC BY 4.0 Open Access | rec: 13 Mar 2018, publ: 13 Mar 2018

Figure 3

Dependence of reaction time and accuracy of task performance on the number of categories, the number of objects, and the spatial frequencies of stimuli.

Left column (“Original”) represents Experiment 1, middle and right columns – Experiment 2 (Low SF – low-pass stimuli, High SF – high-pass stimuli) . Data represent mean values with 95% confidence intervals.

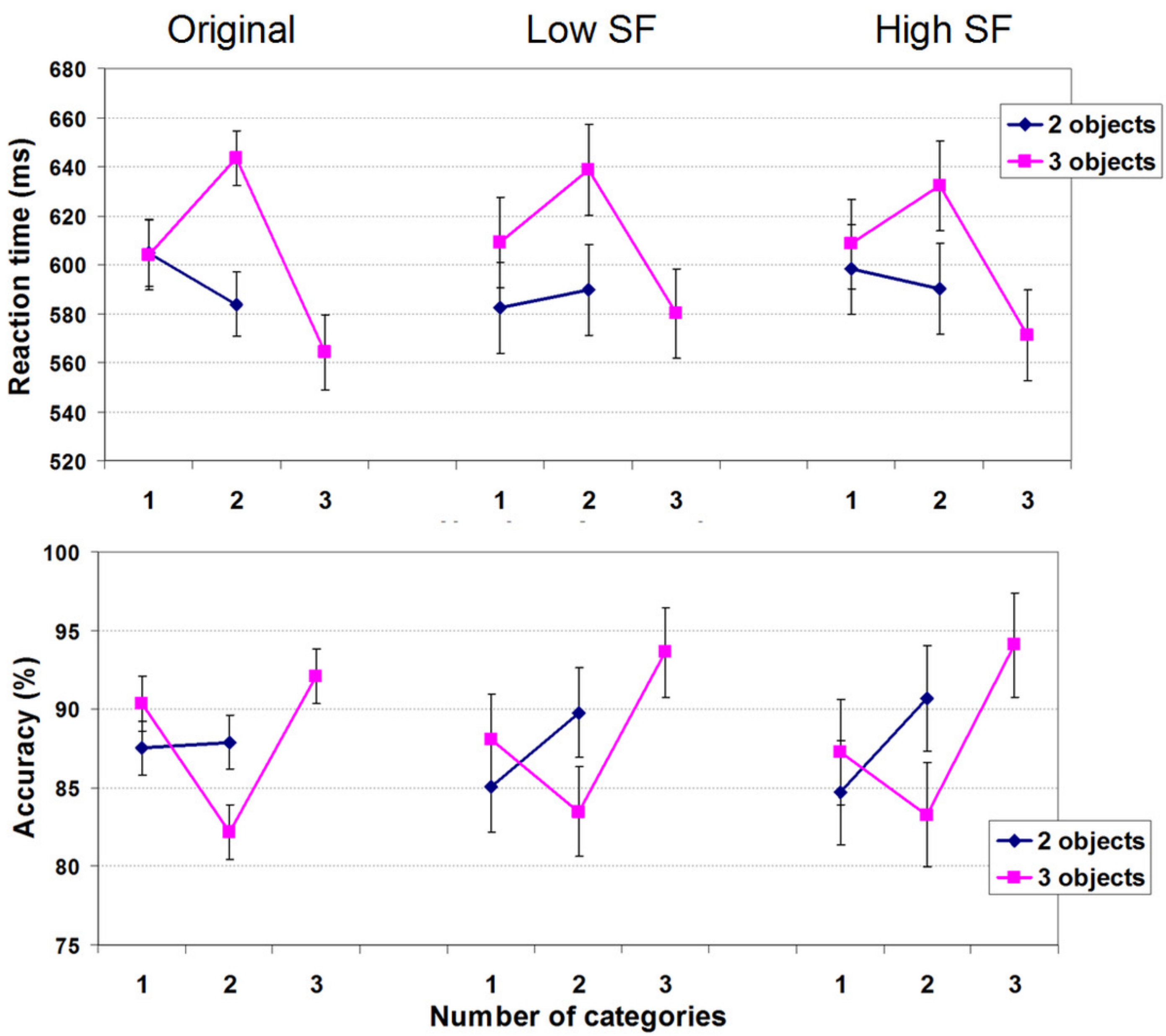


Figure 4

Characteristics of high-pass (solid line) and low-pass (dashed line) filters used for stimulus modification in Experiment 2 and Experiment 4.

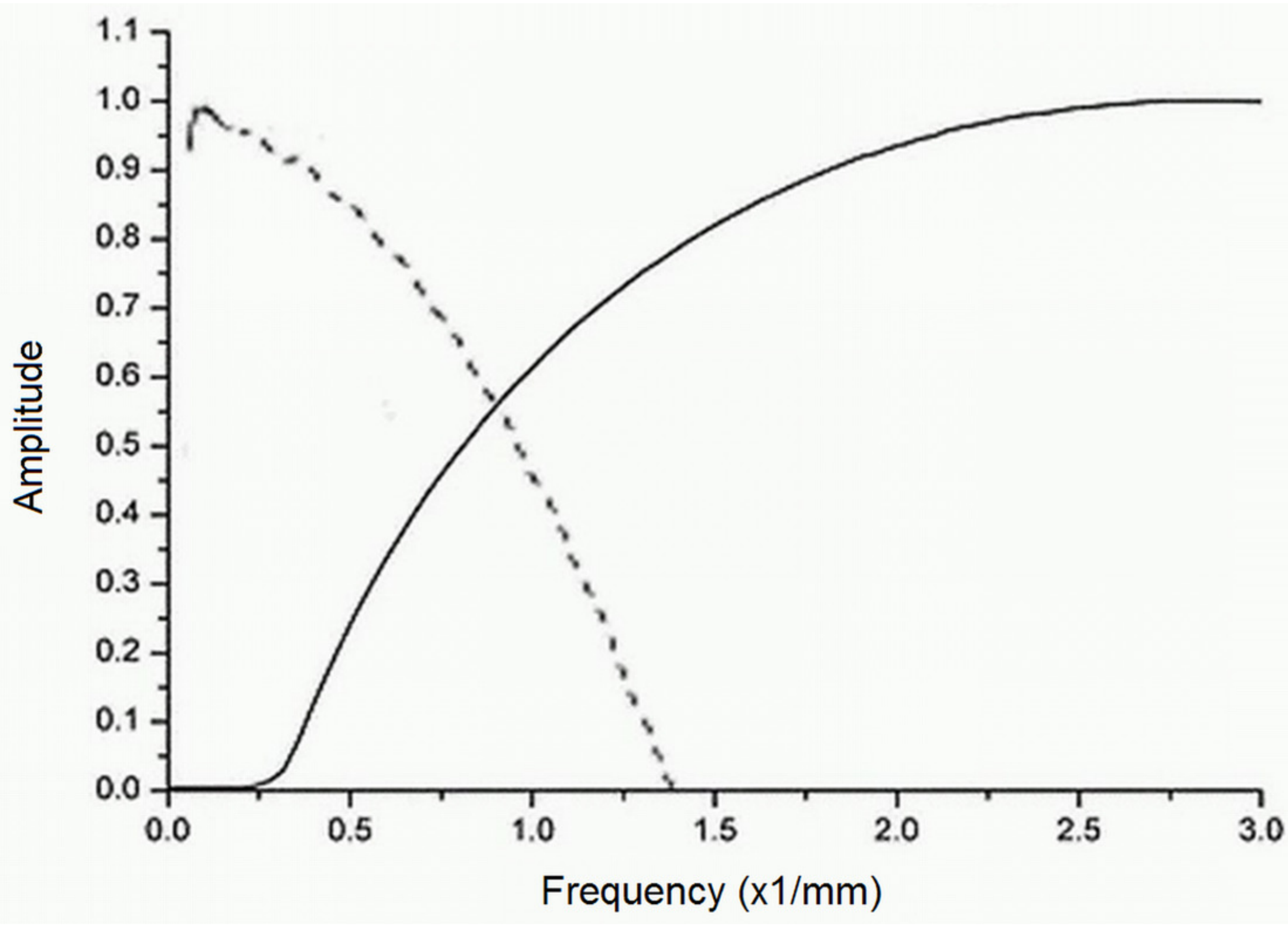


Figure 5

Dependence of reaction time and response accuracy on stimulus type and spatial frequencies of stimuli in Experiment 3 and Experiment 4.

Left column ("Original") represents Experiment 3, middle and right columns – Experiment 4 (Low SF – low-pass stimuli, High SF – high-pass stimuli). Data represent mean values with 95% confidence intervals.

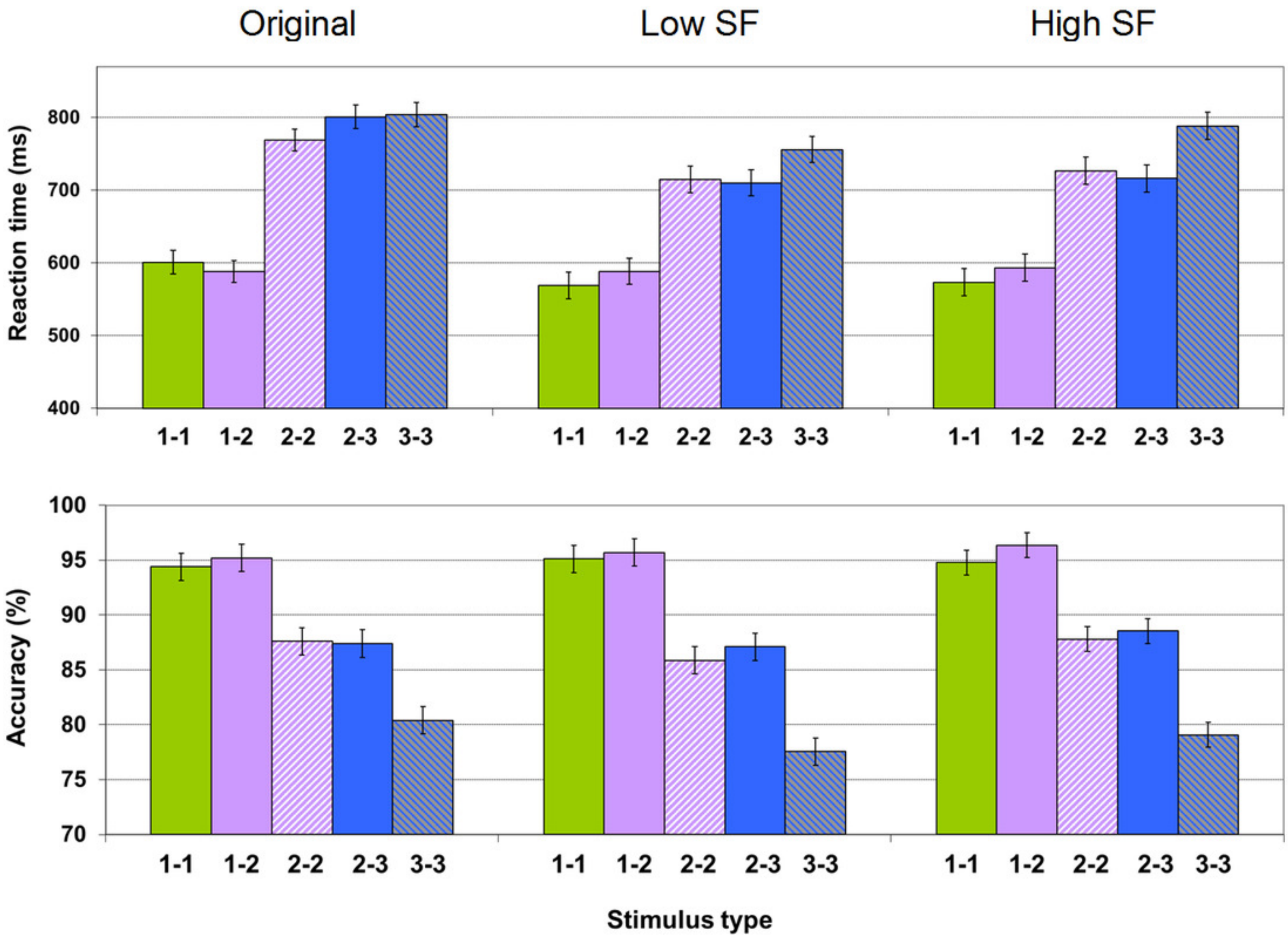


Table 1(on next page)

Significant main effects of Category (number of categories), Objects (number of objects), and their interaction for low-pass and high-pass stimuli on RT and accuracy data.

Factor	Reaction time				Accuracy			
	low-pass stimuli		high-pass stimuli		low-pass stimuli		high-pass stimuli	
	F	p	F	p	F	p	F	p
Category	F(2,94) = 4,07	0.020	F(2,94) = 5.70	< 0.01	F(2,94) = 9,88	< 0.001	F(2,94) = 8.16	< 0.01
Objects	F(1,92) = 5,76	0,018						
Category x Objects					F(1,95) = 10.46	< 0.01	F(1,95) = 8.91	< 0.01