Applied Cognitive Psychology, Appl. Cognit. Psychol. **31**: 109–115 (2017)

More Evidence for Three Types of Cognitive Style: Validating the Object-Spatial Imagery and Verbal Questionnaire Using Eye Tracking when Learning with Texts and Pictures

TIM N. HÖFFLER[1]*, MARTA KOĆ-JANUCHTA[1] and DETLEV LEUTNER[2]

[1]*IPN—Leibniz-Institute for Science and Mathematics Education, University of Kiel, Kiel, Germany*
[2]*University of Duisburg-Essen, Essen, Germany*

Summary: There is some indication that people differ regarding their visual and verbal cognitive style. The Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) assumes a three-dimensional cognitive style model, which distinguishes between object imagery, spatial imagery and verbal dimensions. Using eye tracking as a means to observe actual gaze behaviours when learning with text–picture combinations, the current study aims to validate this three-dimensional assumption by linking the OSIVQ to learning behaviour. The results largely confirm the model in that they show the expected correlations between results on the OSIVQ, visuo-spatial ability and learning behaviour. Distinct differences between object visualizers, spatial visualizers and verbalizers could be demonstrated. © 2016 The Authors Applied Cognitive Psychology Published by John Wiley & Sons Ltd.

INTRODUCTION

Visualizer/verbalizer cognitive style refers to the hypothesis that people differ in their preferences and consistencies in processing visual and verbal information (Blazhenkova & Kozhevnikov, 2009). With reference to dual-coding theory (Paivio, 1978, 1986), which states that two distinct channels are responsible for processing information, people are assumed to have either a visual or verbal cognitive style (e.g. Jonassen & Grabowski, 1993; Mayer & Massa, 2003). That is, people are assumed to think predominantly in either pictures or words (Mayer & Massa, 2003).

Traditionally, many studies and questionnaires that aimed at measuring the visualizer/verbalizer dimension presumed a bipolar distribution of visualizers and verbalizers on the same scale (e.g. Mayer & Massa, 2003; Richardson, 1977). Kirby, Moore, and Schofield (1988) criticized this assumption and emphasized the importance of a more spatial cognitive style. Moreover, the bipolar approach neglected both the plausible assumption of people with hybrid, in-between cognitive styles (cf. Kirschner & van Merriënboer, 2013) and recent results from neuroscience that visual and verbal processing systems are dissociated and anatomically and functionally independent (e.g. Gazzaniga, Ivry, Mangun, & Steven, 2004; Thierry & Price, 2006).

This shortcoming of traditional research in the visualizer/verbalizer dimension might be one of the main reasons why the cognitive style assumption has been criticized quite harshly in the past few years, as the validity of this possibly relevant individual difference has been disputed: For example, Newcombe and Stieff (2012) called the idea, that some learners profit more from visualizations, a myth without sufficient empirical support. In their opinion, ‘almost no evidence that verbalizers and visualizers can be reliably distinguished’ (p. 958) exists. However, other authors disagree; for example, Riding (1997) argued that the construct is valid and that this notion is supported by indices that cognitive style is independent of personality, separate from intelligence and related to, among others, learning performance. Massa and Mayer (2006; also Mayer & Massa, 2003) also found ‘substantial correlations between paper-and-pencil measures of cognitive style and learners’ behaviors during learning’ (p. 334) and concluded that there were indeed visual and verbal learners. What the authors did not find, however, was substantial support for aptitude–treatment interactions (ATIs) of cognitive style and visual versus verbal instructional methods. Pashler, McDaniel, Rohrer, and Bjork (2008) called attention to the lack of well-designed studies investigating such interactions. Indeed, such studies are quite rare, as cognitive style often was investigated only descriptively (Kozhevnikov, 2007, 2013). Massa and Mayer examined the idea that visualizers and verbalizers might profit differently from text or pictures and found only weak support for such an ATI effect, as learning results of visualizers and verbalizers did only differ for very few of the used cognitive style measures.

Additionally, the validity of self-rating questionnaires on cognitive style has been disputed (Green & Schroeder, 1990; Kirschner & van Merriënboer, 2013; Leutner & Plass, 1998). Again, it seems important to show—if they exist—that actual behavioural differences between people who identify themselves can be identified as either visualizers or verbalizers.

Some evidence regarding learning differences can be found in the study of Plass, Chun, Mayer, and Leutner (1998) about multimedia secondary language learning. The combination of textual information with pictures or animations in general resulted in better learning outcome than text alone. But there was a clear ATI effect in that only for visualizers was learning success substantially impaired if pictures or animations were missing (see Leutner & Plass, 1998, for details of the behavioural classification of visualizers and verbalizers). The authors concluded that visualizers profit considerably from visual learning material whereas verbalizers depend far less on visual material. In this study, however, the visualizer/verbalizer dimension was strictly limited to learning preferences. Höffler, Prechtl, and Nerdel (2010) found learning differences on cognitive style between

---

*Correspondence to: Tim N. Höffler, IPN—Leibniz-Institute for Science and Mathematics Education, University of Kiel, Olshausenstr. 62, 24098 Kiel, Germany.
E-mail: hoeffler@ipn.uni-kiel.de

'highly developed visualizers' and 'less developed visualizers' in that the former had better learning results when using static pictures instead of animations while for the latter it made no difference. Riding and Douglas (1993) could also identify an ATI effect when confronting visualizers and verbalizers with either text plus pictures or text plus text: Visualizers performed better with the text-plus-pictures condition, whereas verbalizers were better with the text-plus-text condition.

A more promising approach might be the one by Kozhevnikov and colleagues (Blajenkova, Kozhevnikov, & Motes, 2006; Blazhenkova & Kozhevnikov, 2009; Kozhevnikov, Hegarty, & Mayer, 2002; Kozhevnikov, Kosslyn, & Shephard, 2005): They focused particularly on the development of ecologically valid scales on cognitive style with relations to visual object and spatial abilities and real-world activities. Thereby, they discovered evidence for a further splitting of the visual scale into an object imagery subscale and a spatial imagery subscale. The object imagery scale "assesses preferences for representing and processing colorful, pictorial, and high-resolution images of individual objects and [the] spatial imagery scale […] assesses preferences for representing and processing schematic images, spatial relations amongst objects, and spatial transformations" (Blajenkova et al., 2006, p. 239).

Generally, people with high spatial ability and a spatial visual style score highly on the spatial imagery scale (and are thus called 'spatial visualizers'), while people with an object visual style and low spatial ability score highly on the object imagery scale ('object visualizers'; Kozhevnikov et al., 2002, 2005). A verbal cognitive style is still part of this model. The authors have found some promising results regarding reliability and validity of these three scales, which might lead the way towards the 'hard evidence' for the significance of this individual difference demanded by Pashler et al. (2008).

Most notably, Blazhenkova and Kozhevnikov (2009) showed that people with those three different visualizer/verbalizer cognitive styles chose different college courses in visual arts, physics and writing, respectively (in terms of significant correlations between choice and cognitive style). Likewise, they found a significant relationship between cognitive style and area of specialization: Scientists scored higher than humanities' professionals and visual artists on the spatial imagery scale, while visual artists scored the highest on the object imagery scale and humanities' professionals on the verbal scale.

Thus, there is some evidence that spatial visualizers, object visualizers and verbalizers do exist and that they differ regarding their professional choices. What is lacking, however, is evidence that people with different cognitive styles also show an actual different behaviour when learning—which could serve as an indicator that cognitive style might be a factor to be considered in teaching. The present study aims to close this gap.

Most likely, it could be assumed that cognitive style plays a role in learning with visualizations. Various studies based on Mayer's Generative Theory of Multimedia Learning (e.g. Mayer, 2014) show that the combination of text and pictures usually promotes comprehension and problem-solving transfer (e.g. Plass et al., 1998; Yang, Andre, & Greenbowe, 2003). Mayer's theoretical framework is derived from Paivio's Dual-Coding Theory (Clark & Paivio, 1991; Paivio, 1978). According to this theory, there are two ways of processing information and hence two kinds of mental representations in the cognitive system. In the verbal system, information of a sequential structure such as written text or spoken words is processed. In the non-verbal system, spatial information and pictures are processed. The authors assume that integrating these two cognitive representations properly should improve learning results. It stands to reason that people with different cognitive styles might behave differently when learning with text–picture combinations. While visualizers might focus more on the pictorial information, verbalizers might rely more on the textual information. By linking eye tracking of actual learning behaviour while learning with visualizations to self-assessed cognitive style, the present study aims to validate Kozhevnikov and colleagues' cognitive style dimensions (spatial imagery, object imagery and verbal) and their potential impact on learning and instruction.

MATERIAL AND METHODS

Participants

Thirty-two participants (68.8% female; age $M = 24.63$, $SD = 2.31$ years) from Germany were chosen for this study based on prior telephone interviews with 90 university students of various majors (mostly pedagogics, mathematics, politics and law; science majors and psychology majors were excluded owing to their probable high prior knowledge on the topics we used). These standardized interviews consisted of 14 typical yes/no questions based on the Verbal–Visual Learning Style Rating questionnaire (Mayer & Massa, 2003), the Individual Differences Questionnaire (Paivio & Harshman, 1983) and the Santa Barbara Cognitive Style Questionnaire (Mayer & Massa, 2003). They allowed us to classify participants' cognitive style preliminary as either strongly visual or strongly verbal (if considerably more of the questions were answered in a visual or verbal sense, respectively). Interviewees with neither clear visual nor clear verbal cognitive style were not invited to this study in order to obtain clearly distinguishable, separate groups. This led us to 32 participants, about half of them visualizers and half of them verbalizers according to the preliminary classification. The particular questionnaires were chosen for their relatively easy-to-grasp questions, which also worked on the telephone [which would not have been possible, for example, with the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) by Blazhenkova & Kozhevnikov, 2009].

Measures and questionnaires

First, students' individual cognitive style for visual and verbal material was assessed by a shortened version of the OSIVQ[1] (Blazhenkova & Kozhevnikov, 2009), with 31 items in three subscales, spatial ($\alpha = .86$), verbal ($\alpha = .79$) and object ($\alpha = .93$). The questionnaire had been translated into German. For item examples, please refer to Blazhenkova and Kozhevnikov (2009). Additionally, spatial visualization ability was measured with the paper-folding test (Ekstrom, French, Harman, & Dermen, 1976; Cronbach's $\alpha = .66$), as spatial ability is supposedly linked with spatial imagery cognitive style according to Blajenkova et al. (2006).

Learning tasks and procedure

We chose four text–picture combinations with vastly different topics as stimuli: functioning of the toilet cistern; learned helplessness; Adenosine triphosphate (ATP) synthase; and tying a knot. While the learning material was self-developed, three of its topics were inspired by the works of Hegarty, Kriz, and Cate (2003), Nerdel and Prechtl (2002) and Schwan and Riempp (2004). We chose these particular topics for their wide ranges of domains (e.g. psychology, biology and mechanics), knowledge types (e.g. procedural-motor knowledge and conceptual knowledge) and character (e.g. procedure, process and principle). Thus, we aimed to represent as many typical topics of visualizations as possible to be able to generalize the results and/or to identify differing gaze patterns for different types of pictures. We expected (and checked with four questions prior to the test) participants to have low prior knowledge for each of the topics.

Participants were shown, for each of the four topics, sets of pictures with accompanying texts on a computer screen, each of the sets explaining one of the four topics (Figure 1), while their eye movements were measured with an SMI RED (SensoMotoric Instruments, Teltow, Germany) 120-Hz Eye Tracker, which offered the possibility of free head movements (40 cm × 20 cm at 70-cm distance), an accuracy of 0.4°, a spatial resolution (root mean square) of 0.03° and a sampling rate of 60 and 120 Hz. We decided to present each topic to be learned in a series of three pictures and three boxes of related text. Both sorts of information were placed near to each other (cf. Holsanova, Holmberg, & Holmqvist, 2009; Johnson & Mayer, 2012), and the pictures were designed in a rather abstract manner (i.e. not photo-realistic and without too many details to avoid cognitive overload) in order to facilitate the integration of knowledge. Both texts and pictures were designed to contain comparable information (i.e. were self-contained) in order to enhance the likelihood of choosing one of them as a main source (Chandler & Sweller, 1991; Holsanova et al., 2009; Sweller, van Merriënboer, & Paas, 1998). Furthermore, each set was designed in a comparable way (pictures above, texts below; Figure 1) and was presented for 2.5 minutes (except for the set on learned helplessness, which proved to be much easier to understand and which was therefore presented for 1.5 minutes). The equivalence of text and picture contents was checked beforehand with a small sample of seven participants who were shown pictures or texts separately and asked to explain what they had learned from them. Participants were asked to study the contents as closely as possible, as they had to answer a posttest afterwards.

The posttest included six true/false retention questions for each of the four topics as well as one open question per topic regarding overall comprehension (in case of the topic 'knot-tying', participants were asked to tie the actual knot). Reliabilities ranged from Cronbach's $\alpha = .60$ to .83. Examples are 'The first dog remained on the ground helplessly, because it had been conditioned to expect the electric shots. True or false?'; 'Once the float reaches a certain level, the valve is no longer depressed, which results in new water entering from the pipe. True or false?'; and 'To tie the knot, firstly lay a small loop on the left end. True or false?'.

RESULTS

For analysing the eye-tracking data, each of the four sets of text–picture combinations (representing the four topics to be learned) was analysed with respect to areas of interest (AOIs), that is, regions in the stimuli that we were especially interested in. We created AOIs representing texts and pictures, that is, three AOIs for the three texts, three AOIs for the three pictures and six to eight AOIs for empty space around texts and pictures in every set. For the present analysis, we focused on two typically measured parameters in eye-tracking studies, that is, *dwell time* (sum of durations from all fixations and saccades that hit the AOI in seconds) and *revisits* (number of returns to the AOI after the first visit; Holmqvist et al., 2011). For an easier analysis, an overall composite score of participants' gaze behaviour was generated by adding, in a first step, participants' dwell times and revisits on all pictures and texts (z-standardized) across all stimulus sets, resulting in four different scores. In a second step, these four scores were aggregated into a standardized composite score by calculating the factor scores of their first principal component, which accounted for 71% of the total variance. Table 1 indicates that participants who dwelled longer on pictures and revisited them more often scored positively on this scale, while participants who dwelled longer on texts and revisited them more often scored negatively.

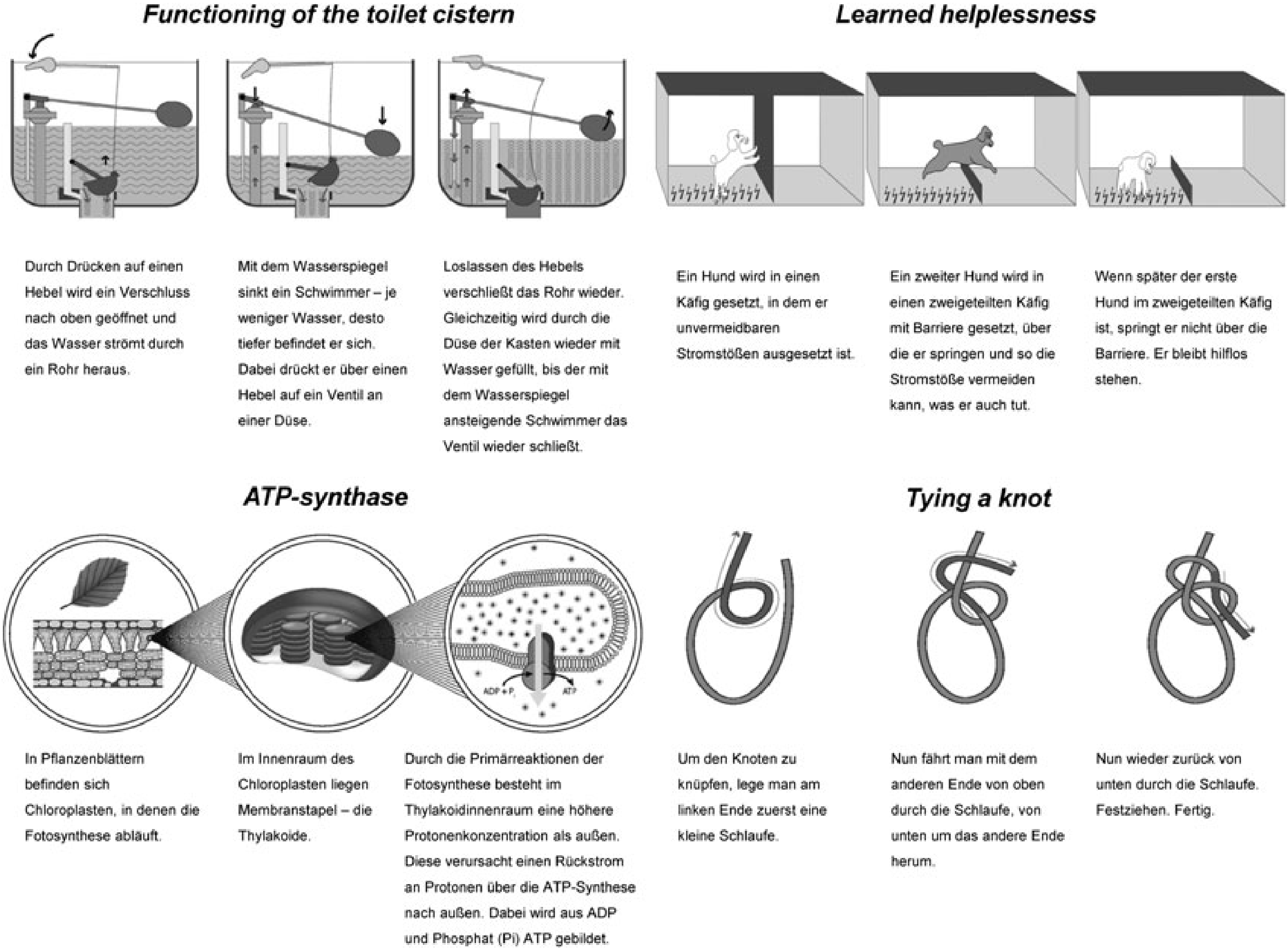
Thus, we created one joint scale for participants' actual gaze behaviours during learning on which participants differed considerably (for a more detailed analysis, see Koć-Januchta, Höffler, Thoma, Prechtl, and Leutner, 2016).

The joint scale was then correlated with scales of cognitive style, visuo-spatial ability and two learning performance scores (retention and comprehension). Table 2 shows the resulting correlations (Pearson and Spearman rho). As can be seen, all three scales of the OSIVQ correlate (unexpectedly) moderately with each other. The verbal scale correlates negatively with both imagery scales.

To get a better impression of the relation between the coefficients, we calculated semi-partial correlations. Figure 2 shows the relationships between a respective

---

[1] We used a shortened version (31 instead of 45 items) because we administered a number of other questionnaires that are not part of the present paper. For the sake of full transparency, they were as follows: the Verbal–Visual Learning Style Rating (Mayer & Massa, 2003), the Individual Differences Questionnaire (Paivio & Harshman, 1983), the Santa Barbara Learning Style Questionnaire (Mayer & Massa, 2003), the Vividness of Visual Imagery Questionnaire (Marks, 1973) and the Verbalizer–Visualizer Questionnaire (Richardson, 1977).



Functioning of the toilet cistern

Durch Drücken auf einen Hebel wird ein Verschluss nach oben geöffnet und das Wasser strömt durch ein Rohr heraus.

Mit dem Wasserspiegel sinkt ein Schwimmer – je weniger Wasser, desto tiefer befindet er sich. Dabei drückt er über einen Hebel auf ein Ventil an einer Düse.

Loslassen des Hebels verschließt das Rohr wieder. Gleichzeitig wird durch die Düse der Kasten wieder mit Wasser gefüllt, bis der mit dem Wasserspiegel ansteigende Schwimmer das Ventil wieder schließt.

Learned helplessness

Ein Hund wird in einen Käfig gesetzt, in dem er unvermeidbaren Stromstößen ausgesetzt ist.

Ein zweiter Hund wird in einen zweigeteilten Käfig mit Barriere gesetzt, über die er springen und so die Stromstöße vermeiden kann, was er auch tut.

Wenn später der erste Hund im zweigeteilten Käfig ist, springt er nicht über die Barriere. Er bleibt hilflos stehen.

ATP-synthase

In Pflanzenblättern befinden sich Chloroplasten, in denen die Fotosynthese abläuft.

Im Innenraum des Chloroplasten liegen Membranstapel – die Thylakoide.

Durch die Primärreaktionen der Fotosynthese besteht im Thylakoidinnenraum eine höhere Protonenkonzentration als außen. Diese verursacht einen Rückstrom an Protonen über die ATP-Synthese nach außen. Dabei wird aus ADP und Phosphat (Pi) ATP gebildet.

Tying a knot

Um den Knoten zu knüpfen, lege man am linken Ende zuerst eine kleine Schlaufe.

Nun fährt man mit dem anderen Ende von oben durch die Schlaufe, von unten um das andere Ende herum.

Nun wieder zurück von unten durch die Schlaufe. Festziehen. Fertig.

Figure 1. The four sets of stimuli used consecutively in the study

Table 1. Loadings of four variables of gaze behaviour on their first principal component

Scale	Dwell time on pictures	Dwell time on texts	Revisits of pictures	Revisits of texts
Factor loading	0.984	−0.983	0.713	−0.635

cognitive style and the two outcomes *after the effects of the other two cognitive styles have been partialled out*.

The object imagery scale correlates significantly ($r = .40$) with participants' gaze behaviour, as does the spatial imagery scale ($r = .46$). However, only the spatial imagery scale additionally correlates with visuo-spatial ability as measured by the paper-folding test ($r = .46$). The verbal scale has only small, non-significant correlations with both gaze behaviour and visuo-spatial ability. Gaze behaviour and visuo-spatial ability also do not correlate significantly ($r = .24$).

For another angle, a *K*-means cluster analysis was calculated in order to check whether verbalizers, spatial visualizers and object visualizers could be clearly identified by their answers on the OSIVQ. The analysis confirmed three separate groups with distinct cluster centres (Table 3),

Table 2. Correlations between the three subscales of the three subscales of the OSIVQ, PFT, and the joint gaze behaviour score when learning with texts and pictures (a positive score indicates focusing on pictures, a negative score indicates focusing on texts)

	OSIVQ_object	OSIVQ_spatial	OSIVQ_verbal	PFT	Gaze behaviour	Retention	Comprehension
OSIVQ_object	1	.522**	−.396*	.088	.594**	−.202	.139
OSIVQ_spatial	.498**	1	−.504**	.458**	.627**	−.113	.206
OSIVQ_verbal	−.413*	−.556**	1	−.342	−.333	−.076	−.074
PFT	.031	.444*	−.287	1	.380*	.081	.502**
Gaze behaviour	.664**	.603**	−.286	.209	1	−.084	.422*
Retention	−.237	−.037	−.061	.236	−.013	1	.053
Comprehension	.137	.178	−.037	.498**	.385*	.384*	1

Note: Both Pearson correlations (in the upper-right triangle) and Spearman rho correlations (in the lower-left triangle) are shown. OSIVQ, Object-Spatial Imagery and Verbal Questionnaire; PFT, Paper-Folding Test. $^{*}p < .05$. $^{**}p < .01$.

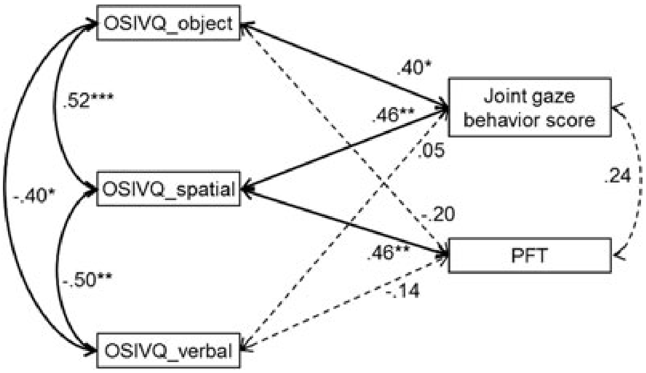


Figure 2. Semi-partial Pearson correlations between the three subscales of the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ), the Paper-Folding Test (PFT) and the joint gaze behaviour score when learning with texts and pictures (a positive score indicates focusing on pictures, a negative score indicates focusing on texts). Effects of the other two OSIVQ subscales are partialled out when looking at the relationship between a style and an outcome. In case of the relation between gaze behaviour and PFT, the effects of all OSIVQ subscales are partialled out. Non-significant correlations are shown as dashed lines. Note that the small sample size prohibited calculating a multivariate path model

Table 3. Cluster centres for three clusters based on participants' ratings on the three scales of the OSIVQ

Scale	Cluster 1 (object visualizers)	Cluster 2 (verbalizers)	Cluster 3 (spatial visualizers)
OSIVQ_spatial	3.43	2.15	3.26
OSIVQ_object	4.12	2.46	2.84
OSIVQ_verbal	3.14	3.98	2.76

Note: OSIVQ, Object-Spatial Imagery and Verbal Questionnaire.

which can justify the participants' classification as object visualizers ($n = 11$), spatial visualizers ($n = 8$) or verbalizers ($n = 13$). All those participants who had been preliminarily classified (in the telephone interview) as visualizers fell into one of both visualizer groups, but some of the previously classified verbalizers fell now into the 'spatial visualizers' cluster. Please note also that those participants in the 'object visualizers' cluster scored highly not only on the object scale (as could be expected) but also on the spatial scale. Thus, the separation between object visualizers and spatial visualizers is not as clear as one might have wished for.

A non-parametric Kruskal–Wallis test then revealed that the three groups differed significantly regarding their gaze behaviour when learning with text–picture combinations, $\chi^2(2) = 19.97$, $p < .001$. To be precise, Mann–Whitney U tests showed that object visualizers ($M = 0.92$, $SD = 0.37$) focused more on pictures than both verbalizers ($M = -0.71$, $SD = 0.98$), $Z = -4.03$, $p < .001$, and spatial visualizers ($M = -0.11$, $SD = 0.54$), $Z = -3.39$, $p = .001$, while spatial visualizers and verbalizers did not differ significantly, $Z = -1.23$, $p = .218$. Figure 3 shows the means and 95% confidence intervals.

Regarding learning differences, composite scores were calculated using z-transformation for each of the topics. Non-parametric Welch tests showed no significant differences between object visualizers, spatial visualizers and verbalizers (Table 4).

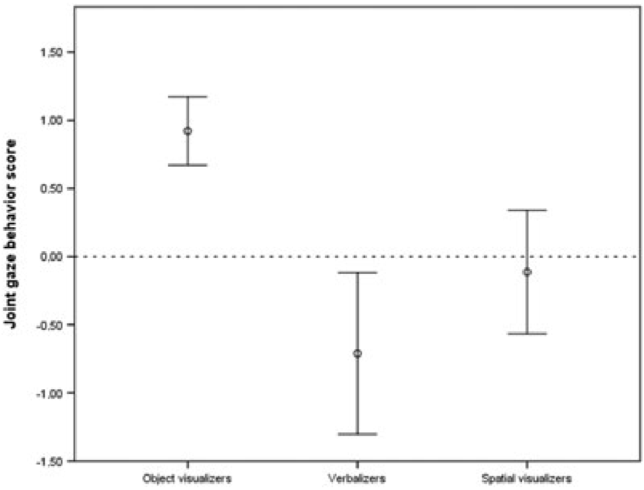


Figure 3. Means and 95% confidence intervals of object visualizers, verbalizers and spatial visualizers on the gaze-behaviour scale when learning with text–picture combinations. Positive values indicate focusing on pictures; negative values indicate focusing on texts

DISCUSSION

The present study aimed to validate Kozhevnikov and colleagues' cognitive style dimensions with regard to actual behaviour during learning with text–picture combinations. Thus, we aimed to assess whether different self-reported cognitive styles, assessed by the OSIVQ (Blazhenkova & Kozhevnikov, 2009), might have actual consequences on learning with visualizations and therefore should be considered when designing learning environments or researching their impact.

Indeed, we could, to a large extent, confirm our expected relations between three different cognitive style dimensions (object imagery, spatial imagery and verbal), visuo-spatial ability and learning behaviour while learning with visualizations. Using eye-tracking methods, we used dwell time on and revisits of pictures versus texts (i.e., aspects of participants' gaze behaviour) as indicators of learning behaviour. We found substantially different correlations of the different cognitive style scales with gaze behaviour and visuo-spatial ability: Apparently, participants scoring high on the object scale and/or the spatial scale of OSIVQ relied more heavily on pictures than on texts (indicated by high positive correlations with a joint gaze behaviour score), while participants scoring high on the verbal scale tended to rely on texts (indicated by a negative, non-significant correlation). Furthermore, only participants scoring high on the spatial scale tended to additionally have a high visuo-spatial ability, as indicated by a significant positive correlation. This is in line with the questionnaire's assumption (Blajenkova et al., 2006; Kozhevnikov et al., 2002) and validates the underlying structure of two types of visualizers and one type of verbalizers. Quite surprisingly, however, we found some indications that spatial visualizers do not focus on pictures as strongly as object visualizers. On the contrary, as indicated by the non-parametric comparisons between three clusters of participants, they did not seem to have a clear preference and tended to use texts and pictures to the same degree. (This apparent contradiction to the correlational results can

Table 4. Results of the Welch tests regarding the different learning outcomes (retention and comprehension) of spatial visualizers, object visualizers and verbalizers for four different text–picture combinations

Learning outcome	Statistic	df1	df2	Sig.
Toilet cistern retention	1.36	2	18.20	.282
Learned helplessness retention	1.91	2	16.35	.180
ATP retention	0.30	2	17.90	.744
Knot retention	3.14	2	15.99	.071
Toilet cistern comprehension	3.29	2	17.31	.061
Learned helplessness comprehension	0.51	2	17.26	.612
ATP comprehension	0.12	2	18.47	.888
Knot tying	0.17	2	17.00	.849

be explained by our finding that participants in the object visualizer cluster scored highly on the object and the spatial scale, while participants in the spatial visualizer scale exclusively scored highly on the spatial scale.) One might speculate that the spatial visualizers' higher spatial ability enables them to process pictures quicker and better and leads to better linking abilities of texts and pictures—which is a superior strategy in any case (e.g. Mayer, 1999). Information from both the verbal and visual channels in working memory needs to be processed and integrated; individual differences in the potential to process visual information (i.e. spatial abilities) on perceptual and cognitive levels (cf. Schnotz, Baadte, Johnson, & Mengelkamp, 2012) might influence retention and understanding of text–picture combinations. Thus, a distinct (object-)visual or verbal cognitive style might also be interpreted in terms of a weakness: Those learners might lack the ability to equally concentrate on both sources of information and are thus not able to integrate verbal and visual information in their working memory. Therefore, spatial visualizers and people with less distinct cognitive style might benefit more from information simultaneously presented in two channels (i.e. the multimedia effect; cf. Mayer, 2014). It might be worthwhile to investigate whether cognitive style and spatial ability indeed influence the process of integrating verbal and pictorial information as described by Schnotz et al. (2012) in a replication study with more participants. In such a study, additional efforts should be made to investigate cognitive style's influence on actual learning outcome, as we did not find any differences. Apparently, while object visualizers, spatial visualizers and verbalizers used quite different strategies to work with text–picture combinations, they all were able to learn sufficiently—even though the significant positive correlation between gaze behaviour and comprehension (Table 2) suggests a relation between a stronger focus on pictures and deeper comprehension. A follow-up study might want to use pictures and texts that are not equivalent in informational content but complement each other.

Moreover, and contrary to expectations based on previous research, the three subscales of the cognitive style questionnaire had medium correlations with each other, which demonstrates related but distinct dimensions. Our correlations were substantially higher than in the original studies (e.g. Blazhenkova & Kozhevnikov, 2009), in which the correlations did not exceed $r = .18$. It stands to reason that these differences could be attributed to our—deliberately chosen—specific sample, which only consisted of people who identified themselves as either clear visualizers or clear verbalizers. With only few or no participants with unclear cognitive styles, these participants tended to mark their self-assessments on both verbal and visual items unambiguously and thus showed rather strong negative correlations between verbal and visual scales. We chose this approach because we expected to obtain clearer results with this specific sample. While we realize that this approach reduces our results' generalizability to a degree (as well as our decision to use a shortened version of the OSIVQ), getting a clear picture whether distinct groups indeed differ on their learning behaviour seemed more important as a first step in our research. In a next step, such a limitation should certainly be removed, and the number of participants should be vastly increased (even though processing large numbers of eye-tracking data is quite time-consuming). Additionally, it would be worthwhile to also assess object ability and verbal ability (cf. Blazhenkova & Kozhevnikov, 2009).

In sum, we could find further evidence that different types of learners can be differentiated according to Kozhevnikov and colleagues' OSIVQ, namely, spatial visualizers, object visualizers and verbalizers. More importantly, we could show that these types of cognitive style are related differently to actual learning behaviour, inasmuch as verbalizers focus their attention on the textual parts of text–picture combinations, while object visualizers focus on the pictorial parts and spatial visualizers do not seem to have a clear preference. Importantly, however, these apparent differences in learning *behaviour* did not seem to have any effects on learning *outcome* in our study. Thus, our results might trigger more research regarding these three types of learners and their apparently crucially different behaviours when learning with texts and pictures. This in turn could lead the way to a systematic consideration of visual/verbal cognitive styles in learning and instruction when designing learning environments or researching their impact, as is already in progress regarding, for example, prior knowledge (Cook, 2006; Kalyuga, 2008) or spatial ability (Höffler, 2010; Höffler, Schmeck, & Opfermann, 2013). Matching the learning environment to people's cognitive style ('matching hypothesis'; cf. Kolb, 1986; Kozhevnikov, Evans, & Kosslyn, 2014; but also Pashler et al., 2008, for an opposite view) might thus enhance learning effects, even though matching material to the specific context seems to be more important and would likely override any cognitive style matching effects (Klein, 2003; Kolloffel, 2012).

On the other hand, if our results regarding a superiority of the spatial visualization strategy should hold true, this might have implications for student instruction in learning. In this case, attempts could be made to encourage and improve this strategy, especially in younger children [that is, matching the styles to the (most common) instructional formats, and not the other way round].

ACKNOWLEDGEMENTS

This research was part of a project funded by the German Research Foundation (grant no. HO 4303/6-1).

REFERENCES

Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-spatial imagery: A new self-report imagery questionnaire. *Applied Cognitive Psychology*, *20*, 239–263.

Blazhenkova, O., & Kozhevnikov, M. (2009). The new object–spatial–verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, *23*, 638–663.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, *90*(6), 1073–1091.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Gazzaniga, M. S., Ivry, R. B., Mangun, G. R., & Steven, M. S. (2004). *The cognitive neurosciences* (3rd edn). Cambridge, MA: MIT Press.

Green, K. E., & Schroeder, D. H. (1990). Psychometric quality of the Verbalizer–Visualizer Questionnaire as a measure of cognitive style. *Psychological Reports*, *66*(3), 939–945.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, *21*, 209–249.

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—A meta-analytic review. *Educational Psychology Review*, *22*(3), 245–269.

Höffler, T. N., Prechtl, H., & Nerdel, C. (2010). The influence of visual cognitive style when learning from instructional animations and static pictures. *Learning and Individual Differences*, *20*, 479–483.

Höffler, T. N., Schmeck, A., & Opfermann, M. (2013). Static and dynamic visual representations: Individual differences in processing. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays: Current perspectives on cognition, learning, and instruction* (edn, pp. 133 – 163). Charlotte, NC: Information Age Publishing.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking. A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology*, *23*, 1215–1226.

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*, 178–191.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction*. Hillsdale: Erlbaum.

Kalyuga, S. (2008). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior*, *24*(3), 852–861.

Kirby, J. R., Moore, P. J., & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology*, *13*(2), 169–184.

Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist*, *48*, 169–183.

Klein, P. D. (2003). Rethinking the multiplicity of cognitive resources and curricular representations: Alternatives to 'learning styles' and 'multiple intelligences'. *Journal of Curriculum Studies*, *35*, 45–81.

Koć-Januchta, M., Höffler, T. N., Thoma, G.-B., Prechtl, H., & Leutner, D. (2016). *Visualizers versus verbalizers: Effects of cognitive style on learning with texts and pictures - an eye-tracking study*. Manuscript submitted for publication.

Kolb, D. A. (1986). *Learning style inventory. Technical manual*. Massachusetts: Institute for Development research.

Kolloffel, B. (2012). Exploring the relation between visualizer–verbalizer cognitive styles and performance with visual or verbal learning material. *Computers & Education*, *58*, 697–706.

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework. *Psychological Bulletin*, *133*, 464–481.

Kozhevnikov, M. (2013). Cognitive style. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology* (edn, pp. 843 – 855). Oxford, UK: Oxford University Press.

Kozhevnikov, M., Evans, C., & Kosslyn, S. M. (2014). Cognitive style as environmentally sensitive individual differences in cognition a modern synthesis and applications in education, business, and management. *Psychological Science in the Public Interest*, *15*(1), 3–33.

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer–verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, *20*, 47–78.

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory and Cognition*, *33*, 710–726.

Leutner, D., & Plass, J. L. (1998). Measuring learning styles with questionnaires versus direct observation of preferential choice behavior: Development of the Visualizer/Verbalizer Behavior Observation Scale (VV-BOS). *Computers in Human Behavior*, *14*, 543–557.

Marks, D. F. (1973). Visual imagery in the recall of pictures. *British Journal of Psychology*, *64*, 17–24.

Massa, L., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer–visualizer cognitive style? *Learning and Individual Differences*, *16*, 321–335.

Mayer, R. E. (1999). Multimedia aids to problem-solving transfer. *International Journal of Educational Research*, *31*(7), 611–623.

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd edn). Cambridge, England: Cambridge University Press.

Mayer, R. E., & Massa, L. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, *95*, 833–841.

Nerdel, C. & Prechtl, H. (2002): *Animation and simulation in biology instruction*. Presentation at the 5th Workshop "Instructional Design for Multimedia Learning" on the European Association for Research on Learning and Instruction (EARLI) SIG 6 Instructional Design, Erfurt University (Germany), 27–29 June 2002.

Newcombe, N. S., & Stieff, M. (2012). Six myths about spatial thinking. *International Journal of Science Education*, *34*(6), 955–971.

Paivio, A. (1978). A dual coding approach to perception and cognition. In H. L. Pick, & E. Saltzman (Eds.), *Modes of perceiving and processing information* (edn, pp. 39 – 52). Hillsdale, NJ: Erlbaum.

Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.

Paivio, A., & Harshman, R. (1983). Factor-Analysis of a Questionnaire on Imagery and Verbal Habits and Skills. *Canadian Journal of Psychology—Revue Canadienne De Psychologie*, *37*, 461–483.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, *9*(3), 105–119.

Plass, J. L., Chun, D., Mayer, R. E., & Leutner, D. (1998). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*, *19*, 211–220.

Richardson, A. (1977). Verbalizer–visualizer: A cognitive style dimension. *Journal of Mental Imagery*, *1*, 109–126.

Riding, R. J. (1997). On the nature of cognitive style. *Educational Psychology*, *17*(1-2), 29–49.

Riding, R. J., & Douglas, G. (1993). The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology*, *63*, 297–307.

Schnotz, W., Baadte, C., Johnson, A., & Mengelkamp, C. (2012). Knowledge acquisition from verbal and pictorial information. In J. Kirby, & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning* (edn, pp. 339 – 365). Cambridge: Cambridge University Press.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, *14*(3), 293–305.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.

Thierry, G., & Price, C. J. (2006). Dissociating verbal and nonverbal conceptual processing in the human brain. *Journal of Cognitive Neuroscience*, *18*, 1018–1028.

Yang, E. M., Andre, T., & Greenbowe, T. Y. (2003). Spatial ability and the impact of visualization/animation on learning electrochemistry. *International Journal of Science Education*, *25*, 329–349.