CrossMark

ORIGINAL RESEARCH

How design guides learning from matrix diagrams

Jan van der Meij[1] · Marije van Amelsvoort[2] ·
Anjo Anjewierden[3]

Abstract Compared to text, diagrams are superior in their ability to structure and summarize information and to show relations between concepts and ideas. Perceptual cues, like arrows, are expected to improve the retention of diagrams by guiding the learner towards important elements or showing a preferred reading sequence. In our experiment, we analyzed how students studied four different versions of informationally equivalent, but differently organized matrix diagrams on personality disorders. The diagram consisted of four header cells on top and four on the left, organizing the content into 16 body cells around four types of personality disorders and four categories within each type. Arrows connected the cells either from left to right or top-down. Participants were 81 university students. We found that the participants mainly studied the diagram based on the four types of personality disorders and that the arrows did not guide initial reading behavior. A posttest with retention questions showed that the participants performed better on the questions that focused on types than on the questions that focused on categories, regardless of the condition they were in. Our study has shown that in matrix diagrams the conceptual

✉ Jan van der Meij
  j.vandermeij@utwente.nl

  Marije van Amelsvoort
  m.a.a.vanamelsvoort@uvt.nl

  Anjo Anjewierden
  a.a.anjewierden@utwente.nl

[1] Faculty of Behavioural, Management and Social sciences, ELAN, Department of Teacher Professional Development, University of Twente, Drienerlolaan 5, P.O. Box 217, 7500 AE Enschede, The Netherlands

[2] Tilburg Institute of Cognition and Communication, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

[3] Faculty of Behavioural, Management and Social sciences, Department of Instructional Technology, University of Twente, Drienerlolaan 5, P.O. Box 217, 7500 AE Enschede, The Netherlands

⏝ Springer

information guided reading behavior more than the orientation of the diagram and perceptual cues.

Keywords Representations · Matrix diagrams · Perceptual cues · Eye-tracking

Introduction

Imagine you are studying psychology. You have to learn four types of personality disorders with four categories of personality traits. These data have been put in a matrix diagram combining the two factors (see Fig. 1). Would the order in which you read the information chunks be guided by the arrows, or would your knowledge of personality disorders or matrix diagrams determine how you read it? Comprehension of visualizations in general is a complex interplay between bottom-up and top-down processes (see Hegarty 2011). In this article, we examine effects of arrows on reading order and recall of information in a matrix diagram. We examine patterns of eye fixations to see how design of the diagram interacts with knowledge, and how reading order influences recall of information.

Matrix diagrams

Diagrams' superiority compared to written descriptions lie in their ability to structure and summarize information, and to make relations between concepts and ideas readily clear (Tversky and Kessell 2006). According to Larkin and Simon (1987), diagrams are advantageous in comparison with—informationally equivalent—written descriptions because they offer a computational benefit. Diagrams aid search and inference processes and allow easier recognition of information. For example, it is easier to see multiple
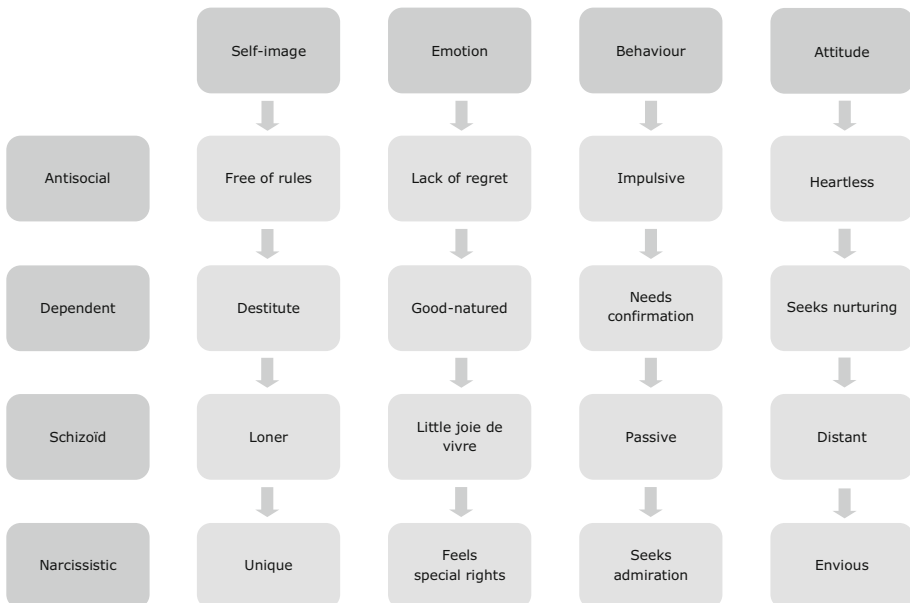

Fig. 1 Matrix diagram on personality disorders

relations between arguments in a two-dimensional diagram than in linear text. Diagrams come in all shapes and sorts, and therefore do not only differ from text, but also differ dramatically from each other, both in appearance and in the way they are processed.

A matrix diagram is a two-dimensional table with headers on top and on the left. It allows topics to be easily compared along one or more categories (Kauffman and Kiewra 2009; Kiewra et al. 1999). Matrices organize by localization, which refers to how close together similar information is placed on a printed page (Larkin and Simon 1987). Kauffman and Kiewra (2009) distinguish between topical and categorical localization. A matrix diagram typically does both, because it combines topics and categories in a two-dimensional structure. For example, our matrix diagram (irrespective of version, see Fig. 2) presents all information about an antisocial personality disorder in adjacent cells (i.e., topical localization) and also presents all information about emotions in different personality disorders in adjacent cells (i.e., categorical localization). Matrix diagrams afford looking for relations between the two dimensions in rows and columns. In a comparison of different types of representations for collaborative learning, Suthers and Hundhausen (2003) found that students working with matrix diagrams searched for relations between cells more than students working with graph or text. Kauffman and Kiewra (2009) found that localization in matrices bolstered both fact (separate cells) and relationship (combinations of cells) learning.
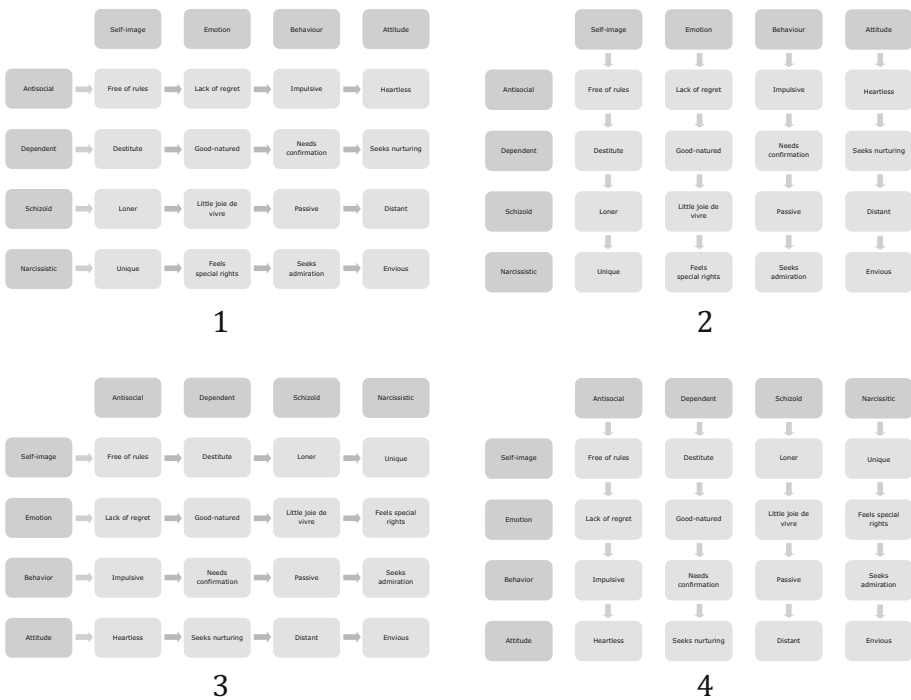

Fig. 2 The four versions of the diagram used in this study (translated)

Processing matrix diagrams

When people process visualizations, both bottom-up and top-down processes play a role. Bottom-up processes include the perception of features in visual objects. These bottom-up processes are intertwined with top-down processes such as attention, driven by the need to accomplish some goal, and with prior knowledge of both the visualization and the content of that visualization (Hegarty 2011). Perceptual cues can thus influence what people attend to, and in turn may influence what they remember from a visualization. For example, Woodman et al. (2003) studied the role of the Gestalt principle of grouping—proximity— in a change-detection task, and found that this bottom-up organization influenced visual memory storage, in that participants performed better if the change was within a perceptual group they paid attention to. The cueing principle, showing people what part of a visualization to attend to, had been extensively studied in multimedia learning (see de Koning et al. 2009; van Gog 2014). Although cueing seems to be successful at guiding attention, it does not always lead to higher learning outcomes (e.g., de Koning et al. 2010). Knowledge may also influence what people attend to. An early account from how readers search for information in a diagram comes from Winn (1993), who started from research of information search in text. He stated that prior knowledge may be more powerful than perceptual guidance of a diagram, since it lets the reader anticipate what to look for, and where to look for it. Studies on comprehension of visualizations, in which experts and novices in a certain domain are compared, show similar results. Experts are able to disregard the salient features if they are not related to the relevant aspects of the visualization (e.g., Haider and Frensch 1999; Lowe 1996).

Our matrix diagram consists of four header cells on top and four on the left, organizing the content into sixteen body cells around four types of personality disorders and four categories within each type. Arrows connect the cells either from left to right or top-down (see Fig. 2). Bottom-up processes in this diagram comprise seeing different units or boxes with information, distinguishing header cells from body cells through different shades of grey, and perceiving the (direction of the) arrows. Top-down processes play an important role in a matrix diagram, first and foremost because it contains a lot of text that must be decoded. In addition, as our participants are told they will receive a test on the content of the diagram, their most probable goal is to remember the diagram. Since we do not ask the participants to focus on specific content and/or elements of the diagram, they won't have expectations about which of them are important. Furthermore, prior knowledge of this type of visualization is needed to understand that headers are distinct from body cells, that a body cell that is placed underneath a header cell is part of that category, and that each body cell is both part of a type of personality disorder and of a category within that disorder. Understanding this can lead to further inferences comparing cells, such as "while an antisocial personality displays impulsive behavior, a schizoid personality displays passive behavior".

Arrows

Arrows are graphical symbols that originate from the pointed shafts used for shooting a bow. The direction is the one from the tail side (the line) to the head side (the triangle). Arrows can be found in texts and graphics, on traffic signs, and even as 21-meter long versions to help airmail pilots find their way across the United States.[1] They are very

---

[1] These concrete arrows were built at the beginning of the 20th century.

powerful cues that are hard to ignore (Heiser and Tversky 2006). In static visuals, they can be used as referential or as attributive markers. Referential arrows refer or point to something in the visual, while attributive arrows contribute to the semantics of the visual (van Wijk et al. 2010), for example by indicating direction (Tversky et al. 2000), or showing causal relations (McCrudden et al. 2007). Because of this myriad of possible meanings, arrows are sometimes misinterpreted by readers of visuals (e.g., McTigue and Flowers 2011; van Wijk et al. 2010).

In our research, the proposed main function of the arrow is an organizational one, intended to guide reading order of different elements in a matrix diagram. Matrix diagrams have no a priori reading order. Nesbit et al. (2007) found that people follow their regular reading behavior when processing node-link diagrams, which may play a role in the inferences people make. This means that people would read a matrix diagram from left to right rather than from right to left, and top to bottom rather than the other way around. However, there is no fixed reading order in terms of rows or columns, although it is important to know what topics belong in one category or which categories can be distinguished within one topic. This makes a matrix diagram suitable for studying the influence of arrows on reading order.

Scale of relational information

In a previous study, we investigated whether arrows in a matrix diagram guided reading behavior, and whether reading behavior affected learning performance (van Amelsvoort et al. 2012). We found that arrows indeed guided reading behavior, that is, participants in our study read more from left to right than top-down if the arrows pointed to the right and read more top-down if the arrows pointed that way. However, content influenced reading order more than the arrows did. We concluded that the difference in *scale of relational information* between the headers might have been accountable for the results (see Tversky et al. 1991).

The relation between information parts is typically described in a range of four scales: nominal, ordinal, interval, and ratio. In statistics, these terms are used to describe relations between variables and the consequences for analyses. In a nominal scale, different parts of information are mutually exclusive categories. In an ordinal scale, the information parts are mutually exclusive, but also rank-ordered. The distances in the ordering are void of meaning. In an interval scale the distances do have meaning and a ratio scale also has an absolute zero. Thus, there is a hierarchy in these scales, in which each level contains the characteristics of the lower level and adds something new (Field 2013; Stevens 1946).

In the previous study (van Amelsvoort et al. 2012), the two header types of the matrix diagram were of a different scale; the headers describing four categories of a baby's development were on a *nominal* level, while the headers describing the month in which the development took place were on an *ordinal*, or even *ratio* level. This means that the month header cells had a logical order from month 1 to month 4, while the categories (physical, intellectual, socio-emotional, and language development) did not. When the categories were placed in the top header cells, the dominant reading order was found to be top-down, while when the categories were placed in the left header cells, the dominant reading order was from left to right (participants' reading order was strongest when the arrows and the category headers both pointed in the same direction). The difference in measurement level may have made it easier to read and recall the months, and therefore may have made the participants focus on the categories.

Indeed, there seems to be a difference between scales in terms of preference and performance. Tversky et al. (1991) found a difference between younger and older children

in representing nominal, ordinal, and interval data, suggesting a difference in difficulty level or perhaps experience with these kinds of representations. Kessell and Tverksy (2011) found that participants, in working with two nominal (place, people) and one ordinal (time) variable in a matrix, preferred to put the ordinal variable along one of the axes rather than in the cells. They found no differences in production, preference, or performance for the location of the nominal versus the ordinal variable on the axes (i.e., no preference for putting the ordinal variable on the horizontal or vertical axis). Thus, there is evidence that there are differences between the scales in terms of preference and performance, although it is unclear whether they would lead to differences in reading order or recall.

Research questions

To test whether the two types of scales were responsible for the results in our previous study (van Amelsvoort et al. 2012), we conducted the present study in which both headers were on a nominal level. We hypothesize that the arrows will exert a greater influence on reading order when both headers present nominal information. If arrows indeed guide reading behavior, the use of arrows can help people to read visualizations in a useful order. Influencing the order in which people read information is of course most interesting when this also influences recall of information. Our research could also show the limitations of using arrows, if we find that people are mainly influenced by content of the headers.

In the current study, we investigated the following research questions:

1. Do perceptual cues (i.e., arrows) and/or content (i.e., nominal headers) guide learners' reading behavior?
2. Does the way in which learners read a matrix diagram influence information recall?

Method

Participants

Eighty-seven participants from the University of Twente in the Netherlands took part in the study for the granting of either money or course credit. The participants were first or second year Behavioral Sciences students, and were randomly assigned to one of four experimental conditions. Four participants were removed from the data set because of technical issues and a further two participants because of an eye tracking percentage lower than two standard deviations of the average. The remaining 81 participants (24 male, 57 female) were between the ages of 18 and 29 years ($M = 21.32$, $SD = 1.99$). The mean percentage tracked was 94.08 ($SD = 3.96$).

Design

A two-by-two between-subjects design was used with diagram orientation (headers types on top or on the left) and cues orientation (arrows from left to right or from top to bottom) as factors, and reading behavior and learning performance as dependent variables. Reading behavior was measured with eye tracking and learning performance was measured with a post-test.

Materials and instruments

Diagrams

The matrix diagram used in this study consists of four header cells on top and four on the left, organizing the content in the sixteen body cells around four types of personality disorders (antisocial, dependent, schizoid and narcissistic) and four categories within each type (self-image, emotion, behavior, and attitude). The four versions of the diagram were informationally equivalent, but differed in orientation. Firstly, the headers containing personality disorders were either displayed on the left while the characteristics were displayed on top, or the personality disorders were on top while the characteristics were on the left. Secondly, the arrows were oriented either from left to right or from top to bottom. Figure 2 shows the four versions of the matrix diagram used.

Questionnaire

Prior to studying the diagram, participants were given a paper-and-pencil demographic questionnaire containing questions on age, gender, education, and academic year. They were also asked to indicate their perceived existing knowledge of personality disorders on a visual analog scale (VAS, indicating a position along a continuous line between two endpoints).

Eye tracker

In this study we used eye tracking to obtain online information on how readers processed the diagram (Hyönä 2010; Mayer 2010; van Gog and Scheiter 2010). Eye tracking informs us about what part of the diagram is attended to and in what order the diagram is read. By using eye tracking, we can find whether learners followed the arrows or not and whether this changes over time. For the eye movement recordings, we used FaceLab 4.5 with two Plea camera's together with Gazetracker version 7. Analyses were done with Gazetracker version 7.

Post-test

For reasons of comparison, we used the same type of post-test as Authors (year), except for the content of the questions. The post-test consisted of twelve questions. Ten questions assessed what participants recalled from the content of the diagram. Five of those, one open question and four multiple-choice questions, required an orientation on the types of personality disorders. The other five questions, one open and four multiple-choice questions, required an orientation on the categories (self-image, emotion, behavior, and attitude). An example of a multiple-choice question is: *Into which category of characteristics falls 'feeling to have special rights'?* with the answer possibilities *self-image*, *emotion*, *behavior* and *attitude*. An example of an open question is: *Describe in note form the characteristics of a dependent personality*. Question eleven showed two versions of the matrix diagram and asked the participants to indicate which one they had studied. In question 12, participants were asked to draw in which order and direction they had studied this diagram.

Interview

In a two-questions interview, the researcher asked the participants to specify how they studied the diagram, and why so. The answers were recorded with a standard digital voice recorder.

Procedure

Each participant was tested individually in a session lasting about 45 min, in a lab of the University of Twente. Two experiment leaders each supervised two of the four experimental groups and alternated the two conditions to get equal amounts of participants in each condition. On arrival, the participant was asked to fill in an informed consent form and the questionnaire. After being familiarized with the eye-tracker and the process in general, the participant underwent eye-tracking calibration. The participant was informed to study one diagram for five minutes and would be given a post-test after the study session to assess recall of the diagram. The matrix diagram was presented for a total duration of 300 s. The participant then answered the questions on the paper post-test, and the questions asked by the experiment leader in the interview.

Analyses

Eye-tracking data

The diagram was divided in 24 look zones, one for each header and body cell (see Fig. 3). For each look zone the following variables were included in the analyses: average fixation duration, cumulative fixation duration, and sequence. *Fixation duration* was defined as the
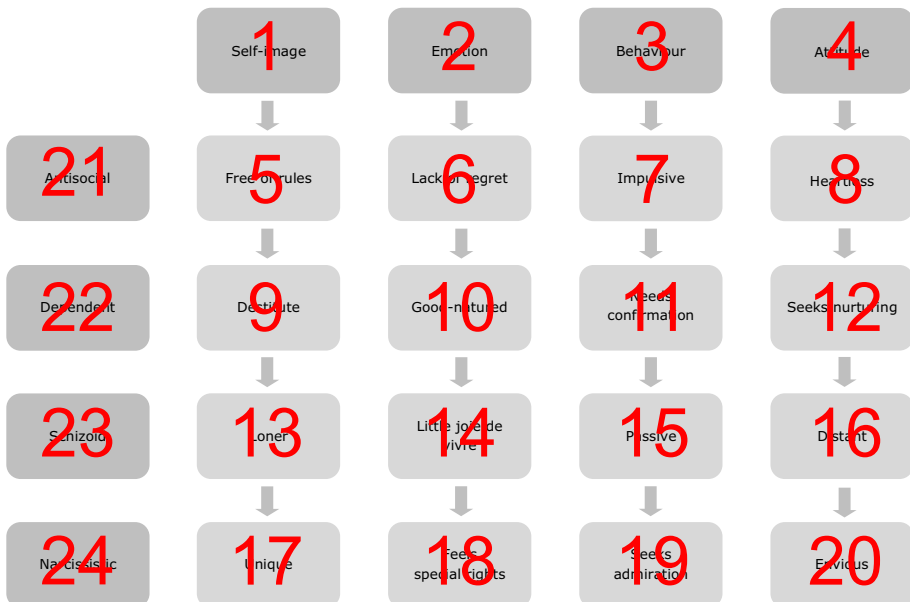

Fig. 3 Definition of look zones

time of relative stability of the eye between two eye movements (saccades). We counted a time of minimal 100 ms within a 30-pixel diameter as one fixation. All fixation durations in a specific look zone (cell) were defined as *cumulative fixation duration*.

We were especially interested in the *sequence*; i.e., the reading order in which the participants studied the diagram. In other words, did the participants study the diagram from top to bottom or from left to right and was the sequence based on perceptual cues? We used the Eye Movement Sequence (EMS) model we developed for an earlier study (van Amelsvoort et al. 2012). Based on the fixations in the look zones, the EMS-model identifies a sequence from left to right (e.g., 5, 6, 7) and top-down (e.g., 5, 9, 13). As long as the time spent in look zones belonging to the sequence are larger than the time spent in look zones not fitting the sequence, the sequence is intact. When the time spent in look zones not fitting the sequence becomes too long, the sequence is terminated and the EMS-model looks for a new sequence. For the analysis of eye movement patterns in the current study, minimally three subsequent row or column look zones had to be present to count as a sequence. The EMS-model takes into account that participants sometimes temporarily deviate from a strict order. For example, when studying the body cells, a participant may temporarily look at a header cell to integrate the information (see Peebles and Cheng 2001). The sequence <5, 6, 7, 8> counts as a strict left-to-right sequence in the same row. The algorithm decides whether <5, 6, 2, 7> counts as a left-to-right sequence depending on the fixation duration in the look zones.

The EMS-model denotes the score of the current sequence after i fixations as S_i and the duration of the ith look zone fixation in the current sequence as d_i. The empty sequence has $S_0 = 0.0$. This gives two possibilities:

1. The next fixation is compatible with the current sequence (e.g., current is <5, 6> and next is <7>). In this case the score is updated with: $S_{i+1} = \max (S_i, d_{i+1})$.
2. The next fixation deviates from the current sequence (e.g., current is <5, 6> and next is <2>). In this case the score is updated with: $S_{i+1} = S_i - C * d_{i+1}$. C is an experimentally validated constant which determines how quickly a sequence is broken if a fixation is not in the correct row or column. In this study C equals 2.0 which means that the fixation duration outside the sequence can be at most half ($1/C$) of the longest fixation duration inside the sequence. If S_{i+1} becomes negative, the sequence terminates.

The algorithm is illustrated using the following example.

Fixation (i)	1	2	3	4	5
Look zone	5	6	2	7	3
Duration (d_i)	0.4	0.5	0.1	0.5	0.3
Score (S_i)	0.4	0.5	0.3	0.5	−0.1

0. Each empty sequence has $S_0 = 0.0$.
1. Look zone 5 is a body cell, so it can start a sequence <5>, $S_1 = \max (S_0, d_1) = 0.4$.
2. Adding 6 results in <5, 6>. A sequence is initiated when two subsequent look zone fixations are adjacent left-to-right (in the same row) or top-to-bottom (in the same column). <5, 6> fits these criteria. The score is updated to $S_2 = \max (S_1, d_2) = 0.5$.
3. Look zone 2 is not in the left-to-right row of the current sequence. The score is updated as follows $S_3 = S_2 - (C * d_3) = 0.3$.

4. Look zone 7 is in the correct left-to-right row: $S_4 = \max (S_3, d_4) = 0.5$.
5. Look zone 3 is not in the row: $S_5 = S_4 - (C * d_5) = -0.1$. The score is now negative and we reject adding 3. The complete sequence is then <5, 6, 2, 7> .

Post-test data

On the eight multiple-choice questions, participants could get one point for every correct answer, thus, receiving a minimum of 0 points and a maximum of 8 points. On the open questions, participants received 1 point for every correct answer that was also in the right position, 0.5 points for a correct answer that was not in the right position or for an incomplete answer in the right position, and 0.25 points for an incomplete answer that was not in the right position, leading to a minimum of 0 points and a maximum of 8 points for both open questions. The maximum total score of the post-test was thus 16 points. As we used a strict response model for the two open questions, the interrater agreement was 100%.

Interview data

The participants' self-reported dominant reading behavior was coded into the direction of reading (left to right, top-down, both left to right and top-down, or random) and recoded to the reading direction according to the diagram they studied (based on the personality disorder types, based on categories, based on both types and categories, or random order).

Results

Do perceptual cues (i.e., arrows) and/or content (i.e., nominal headers) guide learners' reading behavior?

Reading behavior

From the total available time of 300 s, the mean total fixation time in the header cells was 57.42 s (SD = 22.44) and the mean total fixation time in the body cells was 118.96 s (SD 27.29).

To visualize the reading behavior of each of our participants, we created eye movement sequence plots from the EMS-model data. Figure 4 shows part of the plot of five participants. The colored bars show time in sequence in the headers cells or body cells from left to right or top-down. The grey bars show the fixations not indicated as sequence. The black bars show fixations in cells not part of a sequence that are short enough to not break that sequence.

The complete plots of all participants can be found in Appendix 1 in supplementary materials. As can be seen from these plots, the sequences were found across the whole timespan.

We assessed if the arrows and headers influenced reading behavior by performing repeated measures ANOVAs with *time in sequence from left-to-right* and *time in sequence top-down* by *types* and *arrows* for the body cells. Table 1 shows the means and standard deviations for time in sequence per condition.
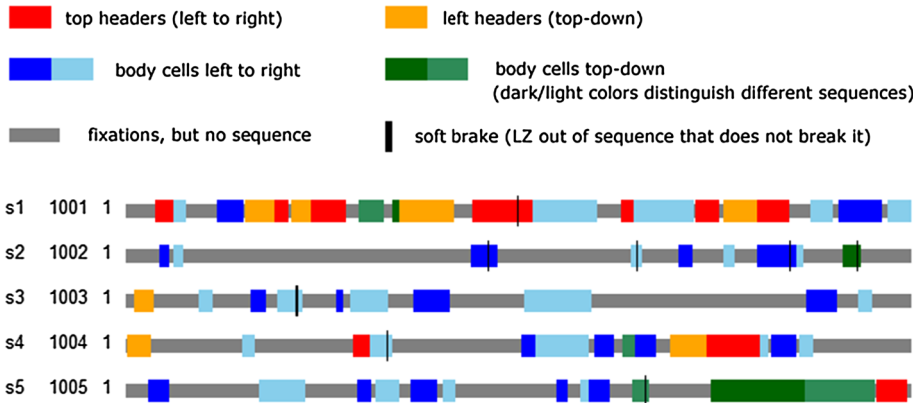
Fig. 4 Part of the eye movement sequence plots of five participants in condition 1 (Types left, Categories top/Arrows left–right)

Table 1 Means and standard deviations for time in sequence per condition in seconds (max. 300 s)

Time in sequence	Condition							
	Types left categories top				Types top categories left			
	Arrows left–right (N = 19)		Arrows top-down (N = 19)		Arrows left–right (N = 22)		Arrows top-down (N = 21)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Left to right	37.41	(26.60)	30.41	(19.90)	16.20	(17.86)	24.28	(25.71)
Top-down	10.06	(10.09)	19.57	(19.61)	30.19	(19.61)	37.75	(25.46)

The repeated measures ANOVAs showed no main effect for reading direction, $F(1,77) = 0.537$, $p = 0.466$. The repeated measures ANOVAs showed an interaction effect between reading direction and the type of header content, $F(1,77) = 20.106$, $p = 0.000$, *partial* $\eta^2 = 0.207$, but no interaction effect between reading direction and the orientation of the arrows, $F(1,77) = 1.195$, $p = 0.278$, *partial* $\eta^2 = 0.015$. The repeated measures ANOVAs showed no interaction effect for reading direction by type of header content and the orientation of the arrows $F(1,77) = 1.355$, $p = 0.248$.

Participant's time in body sequence was longer in the direction of the types than in the direction of the categories, but not longer in the direction of the arrows than in the direction without arrows. Figure 5 shows the participant's time in body sequence per condition in seconds. When the types are presented in the left headers, the overall left to right sequences are longer compared to the top-down sequences. When the types are presented in top headers, the overall top-down sequences are longer compared to the left to right sequences.
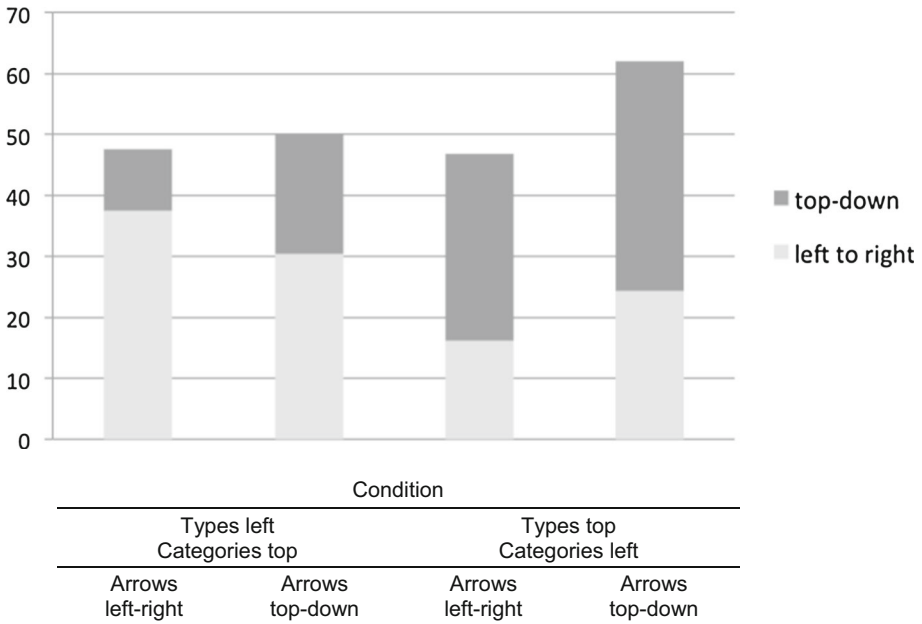
Fig. 5 Participant's time in body sequence per condition in seconds (max. 300 s)

Table 2 Initial sequence per condition

Initial sequence	Condition			
	Types left categories top		Types top categories left	
	Arrows left–right N	Arrows top-down N	Arrows left–right N	Arrows top-down N
Left headers	11	8	15	13
Top headers	4	3	1	5
Body cells left to right	3	5	4	0
Body cells top-down	1	3	2	3
Total	19	19	22	21

Initial sequence

To examine how the participants approached the diagram, we analyzed the initial sequence. Overall, for 58% of the participants the initial sequence was in the left headers, 16% in the top headers, 14.8% in the body cells from left to right, and 11.1% in the body cells top-down. Table 2 shown the initial sequence of the participants per condition.

To find out whether participants initially followed the perceptual cues, we analyzed the direction of the initial sequence in parsing the body cells. This showed that in the diagrams in which arrows pointed from left to right, 51.2% of all participants' initial reading sequence was also from left to right, and in the diagrams in which arrows pointed top–down, 61.5% of all participants' initial reading sequence was also top–down, $\chi^2(1) = 1.31$, $p = 0.252$, $\lambda = 0.067$. This result shows that the arrows did not influence initial reading

Table 3 Initial sequence body cells (did they follow the cues?)

Initial sequence body cells	Condition			
	Types left categories top		Types top categories left	
	Arrows left–right N	Arrows top-down N	Arrows left–right N	Arrows top-down N
Left to right	16	10	5	5
Top-down	3	8	17	16
Total	19	18*	22	21

*For one participant, no sequence found by the EMS-model

Table 4 Self-reported dominant reading behavior

Dominant reading behavior	Condition			
	Types left categories top		Types top categories left	
	Arrows left–right N	Arrows top-down N	Arrows left–right N	Arrows top-down N
Types	14	14	13	10
Categories	3	3	3	7
Both	1	1	5	4
Random	1	1	1	0
Total	19	19	22	21

behavior. Table 3 shows the initial sequence in the body cells of the participants per condition.

Self-reported dominant reading behavior

Analyses of the interview questions showed that overall 63% of the participants self-reported having studied the diagram primarily based on the personality disorder types, 19.8% based on categories, 13.6% based on both types and categories, and 3.7% in random order. Table 4 shows the self-reported dominant reading behavior of the participants per condition.

Does the way in which learners read a matrix diagram influence information recall?

Overall, participants answered about half of the post-test questions correctly, with a mean of 8.59 (SD = 2.71) out of 16 points.

To compare the learning results of the four experimental conditions, a MANOVA was conducted with type of header content and the orientation of the arrows as independent variables, and with scores on questions that focus on types and scores questions that focus on categories as dependent variables. The MANOVA showed a trend of type of header content, Wilks's Lambda = 0.932, $F(2,76) = 2.756$, $p = 0.070$, no effect of orientation of the arrows, Wilks's Lambda = 0.996, $F(2,76) = 0.157$, $p = 0.855$, and no interaction effect of types and arrows, Wilks's Lambda = 0.992, $F(2,76) = 0.314$, $p = 0.732$.

Univariate results for type of header content showed a significant effect on questions that focus on types, but not on questions that focus on categories, $F(1,77) = 5.549$, $p = 0.021$ and $F(1,77) = 0.972$, $p = 0.327$. Univariate results for orientation of the arrows content showed no significant effect on questions that focus on types, nor on questions that focus on categories, $F(1,77) = 0.168$, $p = 0.683$ and $F(1,77) = 0.019$, $p = 0.889$. Univariate results for type of header content and orientation of the arrows showed no significant effect on questions that focus on types, nor on questions that focus on categories, $F(1,77) = 0.025$, $p = 0.874$ and $F(1,77) = 0.577$, $p = 0.450$.

Conclusion and discussion

Irrespective of the experimental condition, we found a dominant reading order based on the content of the diagram. This was true both for the overall reading order and for the initial sequence.

The focus on content confirms the importance of top-down processes in processing diagrams. Top-down processes consist of attention given to certain areas of the diagram which is driven by goals, and prior knowledge of both the visualization and its content (Hegarty 2011). Participants' account of why they studied the diagram the way they did, implies that they were driven by the content of the diagrams. Although our diagram contained two nominal scales, most participants took the types of personality disorders as more important than the categories. The types of personality disorder seem of a *higher* level than the categories within each type. Therefore, it seems 'most logical' to follow the types, as some participants stated.

Also, prior knowledge of the visualization itself may have played a role. Matrices are known for affording comparison between categories by relating two types of knowledge (given in header cells) for each body cell. Because of the nature of matrix diagrams, the arrows, intended to guide the participants' reading behavior, were not needed for understanding the diagram. They seemed to indicate a special relation between the boxes connected through arrows, while in fact there was no such special relation. This can be seen as an unwanted side-effect of the many meanings arrows are able to serve (Tversky et al. 2000). It appears that in the case of matrix diagrams, arrows as cueing devices seem not to be of special relevance with respect to organizational guidance. They are overruled by the conventional format and by the content presented in the diagram. Participants may have tried to assign meaning to the arrows, but may have neglected them when no meaning could be inferred. As may be expected, most participants first looked at the header cells before looking at the body cells. Header cells provide a framework in which the content of the body cells can be placed. The interaction of the diagrams' conventions and the readers' knowledge (Winn 1993) probably led the participants to encode the information in the matrix diagram based on types. The two-dimensional organization of the matrix diagram seems so powerful that the additional inclusion of arrows as organizational cues seems either futile or even having detrimental effects.

It is, however, remarkable that even the initial sequence was not led by arrows. If arrows are very powerful cues that are hard to ignore (Heiser and Tversky 2006), and are related to bottom-up perceptual processes (Ware 2012), we would expect participants to follow arrows at least initially. A possible explanation for this finding is that there were no arrows in the headers, and people seemed to have read the headers first. The headers also stand out perceptually, because they are grey. The grey headers instantaneously show that the

diagram is a matrix diagram, which may make the arrows less important. In our previous study (Authors, year), we did find an effect of arrows on initial sequence and also on general reading order (though content played a more important role than arrows did). The diagram in our earlier study contained a nominal (types of development) and ordinal level (the first four months of a baby's life), while the current study contained two nominal levels. It seems that studying content at an ordinal level is influenced more by arrows than studying content at a nominal level. This idea is corroborated by the fact that the time-in-sequence was less than half compared to our earlier study, while both had the same overall study time. In other words, participants in our current study far less read three cells in a row (either from left-to-right or top-to-bottom). This makes sense, because there was no order between adjacent cells. In summary, the question whether arrows in matrix diagrams guide reading behavior can be negated. Although *some* participants followed the arrows in the beginning, most of them switched when the arrows did not follow the type of personality disorders.

The question whether reading order influences learning performance is tricky to answer. The results showed that only the type of header content had an influence on learning. Most participants followed the types of personality disorders, and questions on types of personality orders have more often been answered correctly. Condition had no influence on learning performance. They did not differ on answering the test questions in general, nor in the type and category questions in particular. Thus, whether the arrows pointed from left to right or from top to bottom made no difference for learning. Although the organization of information has an important influence on how it is stored in memory (Woodman et al. 2003), we cannot infer that there is a causal relation between reading order and correct answers. It is plausible that the content is more important than the reading order. The types of personality *can be distinguished* on several categories, rather than the categories *defining* the types.

Our results entail that even with such pre-structured learning material like the one used in this study, readers work actively with the presented information. Instead of simply following the implemented design elements, they make combinations and compare and anchor information according to *their* main point of reference. This should by no means discourage designers of learning materials but should simply remind them of the importance of the input of learners and direct the focus to the needs of the learners and the task they are to perform. Diagrams should be designed to facilitate the use of effective strategies of encoding, the process whereby sensory information is transformed into a representation suitable for storage in memory and subsequent retrieval (Winn 1990). What seems to be crucial for adhering to cues is their function and the content to be presented. Rather than arbitrarily grafting cues onto diagrams, the use of cues should be tailored to the specific processing demands the diagram poses (Lowe and Boucheix 2010). In the case of matrix diagrams, the use of arrows should be considered with care, since the matrix diagram is in itself a good way to organize relations between two variables.

The idea that arrows have no added value in matrices is also the most important limitation of our study. We deliberately chose matrix diagrams for our study, because there is no a priori reading order or reason to use arrows, and because arrows can easily be manipulated in both directions of the matrix diagram. To answer the question whether arrows will influence reading order in visualizations, and whether this reading order in turns influences learning performance, future studies should look at other types of visualizations in which arrows play a natural role. In a future study we will therefore manipulate arrow use in argument diagrams. In argument diagrams, arrows indicate (causal) relations between arguments, or between arguments and conclusions. Arrows from

arguments in favor to arguments against may for example lead to different viewpoints than arrows from arguments against to arguments in favor, even though the information is the same. This way, we can investigate whether direction of arrows leads to different interpretations of the same information.

References

de Koning, B. B., Paas, F., Tabbers, H. K., & Rikers, R. M. J. P. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*(2), 111–122. doi:10.1016/j.learninstruc.2009.02.010.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113–140. doi:10.1007/s10648-009-9098-7.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: Sage.

Haider, H., & Frensch, P. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25*(1), 172–190. doi:10.1037/0278-7393.25.1.172.

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science, 3*(3), 446–474. doi:10.1111/j.1756-8765.2011.01150.x.

Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science, 30*(3), 581–592.

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*(2), 172–176. doi:10.1016/j.learninstruc.2009.02.013.

Kauffman, D. F., & Kiewra, K. A. (2009). What makes a matrix so effective? An empirical test of the relative benefits of signaling, extraction, and localization. *Instructional Science, 38*(6), 679–705. doi:10.1007/s11251-009-9095-8.

Kessell, A., & Tversky, B. (2011). Visualizing space, time, and agents: Production, performance, and preference. *Cognitive Processing, 12*(1), 43–52. doi:10.1007/s10339-010-0379-3.

Kiewra, K. A., Kauffman, D. F., Robinson, D. H., Dubois, N. F., & Staley, R. K. (1999). Supplementing floundering text with adjunct displays. *Instructional Science, 27*(5), 373–401. doi:10.1007/BF00892032.

Larkin, J. H., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 99,* 65–99.

Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education, 11*(4), 377–397. doi:10.1007/BF03173279.

Lowe, R. K., & Boucheix, J.-M. (2010). Manipulatable models for investigating processing of dynamic diagrams. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), *Diagrammatic Representation and Inference: 6th International Conference, Diagrams 2010, Portland, OR, USA, August 9–11, 2010. Proceedings* (pp. 319–321). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-14600-8_41

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction, 20*(2), 167–171. doi:10.1016/j.learninstruc.2009.02.012.

McCrudden, M. T., Schraw, G., Lehman, S., & Poliquin, A. (2007). The effect of causal diagrams on text learning. *Contemporary Educational Psychology, 32*(3), 367–388. doi:10.1016/j.cedpsych.2005.11.002.

McTigue, E. M., & Flowers, A. C. (2011). Science visual literacy: Learners' perceptions and knowledge of diagrams. *The Reading Teacher, 64*(8), 578–589. doi:10.1598/RT.64.8.3.

Nesbit, J. C., Larios, H., & Adesope, O. O. (2007). How students read concept maps: A study of eye movements. *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 961–3970. Retrieved from http://summit.sfu.ca/system/files/iritems1/12811/etd6060_KBisra.pdf

Peebles, D., & Cheng, P. C. H. (2001). Graph-based reasoning: From task analysis to cognitive explanation. In *Paper presented at the Twenty-third Annual Conference of the Cognitive Science Society, Edinburgh*.

Stevens, S. S. (1946). On the theory of scales of measurement author (s): S S Stevens. *Science, 103*(2684), 677–680. doi:10.1126/science.103.2684.677.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *The Journal of the Learning Sciences, 12*(2), 183–218. doi:10.1207/S15327809JLS1202.

Tversky, B., & Kessell, A. (2006). Using gestures and diagrams. In *Earli SIG 2 bi-annual meeting "text and graphics comprehension"* (pp. 8–9). Nottingham. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.5064&rep=rep1&type=pdf#page=12

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23*(4), 515–557. doi:10.1016/0010-0285(91)90005-9.

Tversky, B., Zacks, J., Lee, P., & Heiser, J. (2000). Lines, blobs, crosses and arrows: Diagrammatic communication with schematic figures. In M. Anderson, P. Cheng, & V. Haarslev (Eds.), *Theory and Application of Diagrams* (pp. 221–230). Berlin, Heidelberg: Springer. http://doi.org/10.1007/3-540-44590-0_21

van Amelsvoort, M., van der Meij, J., Anjewierden, A., & van der Meij, H. (2012). The importance of design in learning from node-link diagrams. *Instructional Science, 41*(5), 1–15. doi:10.1007/s11251-012-9258-x.

van Gog, T. (2014). The signaling (or cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2, revised ed., pp. 263–278). New York: Cambridge University Press.

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*(2), 95–99. doi:10.1016/j.learninstruc.2009.02.009.

van Wijk, C., Speessen, R., & Maes, A. (2010). Interpreting arrows in static pictures. In *Interpreting Arrows in Static Pictures*. Paper presented at the EARLI SIG 2 Meeting, Tubingen.

Ware, C. (2012). *Information visualization: Perception for design* (3rd ed.). Waltham, USA: Elsevier.

Winn, W. (1990). Encoding and retrieval of information in maps and diagrams. *IEEE Transactions on Professional Communication, 33*(3), 103–107. doi:10.1109/47.59083.

Winn, W. (1993). An account of how readers search for information in diagrams. *Contemporary Educational Psychology, 18*(2), 162–185. doi:10.1006/ceps.1993.1016.

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review, 10*(1), 80–87. doi:10.3758/BF03196470.