A smaller scope: How can we get them to remember a specific piece of information about a graph?

We want to make a piece of information about the graph more salient.

1. We *prime* the audience by telling them that a piece of information is important, or should be paid attention to.
2. We *frame* that information in terms of the shape surrounding it. We mention points or parts of the shape that are near it or that include it.
3. We make that frame more memorable by prioritizing parts that are abnormal; the parts that defy expectations for the shape.

For this test, we try to get users to remember the peak average value (that we tell them). We give them several scatterplots, where it may be difficult to determine the information of interest using visuals alone.

Case 1 (control), user sees each graph, reads the peak average value, waits some time, then tries to recall the maximum value for each graph. 5 graphs total.

Case 2, user sees each graph, reads its description (which includes the peak average value), waits some time, then tries to recall the maximum value for each graph. 5 graphs total.

How can we generate a description about a visual graph to help the reader remember a specific piece of information about it? More specifically, can we help the reader remember the value of a specific point?

As a specific task, we use a single piece of information that is calculable for all graphs, but is difficult to determine by visuals alone in a scatter plot: the peak average value and when it occurs. We call this the *point of interest*.

In this solution, we attempt to make the point of interest more salient using three things: priming, framing, and abnormality.

1. We *prime* the audience by telling them that a piece of information is important, or should be paid attention to, and where to expect it, before revealing the information itself later on. Priming, as an effect, is well known, but in the context of story, mentioning the information but revealing it later also may increase interest, and gives the reader both something to look for and a lens through which to interpret what they're about to read.
2. We *frame* that information in terms of a shape in the graph, such as a 'w', a 'v', or an upward or downward line. We mention *critical points* of that shape that are near to, or include, the point of interest. Critical points are those points which identify the shape; a concept motivated by early work on shape description. Framing, too, as an effect, is well known. In the context of this task, by relating the point of interest to an identifiable shape and its critical points, we give it a spatial frame in the graph. 🡨 Sounds natural, but harder to prove experimentally
3. We make that frame more memorable by describing how it is *abnormal*. We mention parts of the shape that are near the point of interest or that include the point of interest which defy some of the shape's normal set of values, such as asymmetry or whether its dimensions are at a good ratio. Doing so makes descriptions of otherwise repetitive shapes unique, and, in the context of story, parts that defy expectations are sometimes more easy to remember.

To test the effectiveness of this method, we wish to see whether, when given a description as generated above, subjects will more often remember the x and y value of the point of interest.

The procedure is as such:

1. The user looks at a random series of 5 graphs and their descriptions. Each graph is a scatter plot of information gathered from Lake George about the levels of certain chemicals at specific sites. The chemical and site name are changed to letters (e.g. Chemical A at Site F).
   1. In Case 1, the control, the description consists only of the x and y value of the point of interest.
   2. In Case 2, the description consists of one generated by the program, which includes the x and y value of the point of interest.
2. After a period of waiting, the graphs are shown again to the user and the user is asked to give the peak average value and when it occurs for each graph.
3. The performance of each user is measured by the error between the value and date they give for each graph and the value and date that was given in the graph's description.

* Y value digits may be hard to remember
* And dates may need to be referred to the x axis tick marks
* Some of the sentences, it’s not obvious what the user should be looking at
* The interjections (like however and but) suggest that it’s different from the previous trend
* Sentences don’t read terribly fluid (especially the information reveal and the references in the V shape). Should also be less ambiguous (especially with the V shape).

Do several tests:

* Make minor adjustments, run tests over
* Design several experiments

We’re claiming that all of the techniques that go into the paragraph help people with remembering the information, and some of it goes into describing what’s wrong with the shape.

* Test the abnormality part separately
* See if it helps them remember the shape
* Change the points a bit so the scatter plot looks better; then, when testing, have the imperfect image paired with the perfect (or otherwise altered image).
  + In the experimental condition, have the description paired with the abnormal graph
  + In the control, there is no description
  + Afterwards, we show them both the abnormal graph ^ and a fixed or altered graph of the same shape, and ask them which one is the graph they have seen.
* This will help with the abnormality case.

Instead of asking for the y value, ask for the year.

* Or instead of asking for the specific value, ask them to choose the correct one

Also, the paragraph is long; it probably doesn’t work better than just presenting the information sentence by itself!

* So, try with the difference between two cases being whether or not the priming sentence is present.
* Try with priming sentence at the very beginning? It doesn’t flow well where it is right now.
  + Or add another sentence right after priming telling them we’re going back to the regular description.
  + Rewrite the priming sentence, and either leave it in the second sentence or bring it to the first sentence, and make it clear that the information that we’re hinting at isn’t right after the priming sentence.
* The year we’re referring to should be on the graph; it’s a memory test, and it could unnecessarily complicate the situation

What if we brought it to a different dataset?

* Like stock prices?
* Look into this later

Description Generation:

1. The graph as a whole is segmented into a PLA.
2. Shapes in the graph are identified. Currently, the boundaries for each shape are placed by hand. The shapes that can be identified are a ‘w’, a ‘v’, and a line. For each shape, a set of critical points are identified with it.
   1. For a ‘w’, the critical points are:
      1. The starting peak
      2. The first trough
      3. The middle peak
      4. The second trough
      5. The ending peak
   2. For a ‘v’, the critical points are:
      1. The starting peak
      2. The middle trough
      3. The ending peak
   3. For a line, the critical points are:
      1. The starting point
      2. The ending point
3. For each shape, a set of ‘normal’ values are calculated for the x and y value of each critical point. What these normal values are is dependent on the shape.
   1. For a ‘w’, the normal values are based off of two things:
      1. Symmetry; the shape should be symmetric about the middle x point:
         1. The starting peak should have the same y value as the ending peak.
         2. The first trough should have the same y value as the second trough.
         3. The middle peak x value should be halfway between the starting x value and the ending x value.
      2. Ratio; the dimensions of the shape should be square.
         1. The distance between the topmost y value and the bottommost y value should be equal to the distance between the leftmost x value and the rightmost x value.
   2. For a ‘v’, the normal values are based off of two things:
      1. Symmetry; the shape should be symmetric about the middle x point:
         1. The starting peak should have the same y value as the ending peak.
         2. The middle trough x value should be halfway between the starting x value and the ending x value.
      2. Ratio; the dimensions of the shape should be square.
         1. The distance between the topmost y value and the bottommost y value should be equal to the distance between the leftmost x value and the rightmost x value.
   3. For a ‘line’, the normal values are based off of three things:
      1. Slope Direction; the lines should all be going either up or down.
      2. Slope Magnitude; between any adjacent segments going in the same direction, the magnitude of their slopes should be equal.
      3. Slope Ratio; the overall slope should be 1:1 with the x-y ratio of the graph as a whole.
4. Any points that deviate from the normal values above are labelled as ‘abnormal,’ and a comment on why they are abnormal is generated.
5. A hint on where the point of interest may be located is generated relative to the shape. If any points on the shape are mentioned at this step, they are counted as related points.
   1. On all shapes, the point of interest is checked against all critical points. If it matches one of the critical points, then the matching critical point is the only related point.
   2. On a ‘w’ or a ‘v’, if the point of interest falls between two critical points, those points are counted as related points.
   3. On a ‘line,’ if the point of interest does not fall on a critical point, its location is referred to as a fractional position on the line (e.g. halfway through, a quarter of the way through, etc.)
6. A subset of abnormalities which are closest to the point of interest and its related points are chosen.
7. A subset of critical points which are closest to the point of interest and its related points, as well as those which are close to abnormalities, are chosen. This allows the characterization of the shape to be tailored towards around the abnormalities that will be presented in the future, the related points that will be presented in the future, and the point of interest itself.
8. The shape as a whole is given a shallow summary, stating what the shape is and when it starts and ends. If one of the start or end points is already included as a critical point, it is not mentioned here to prevent repetition.
9. A description revealing the point of interest itself is generated.
10. The order of presentation is as follows:
    1. Shallow summary (8.)
    2. Point of interest hint (5.)
    3. Critical point presentation (7.)
    4. Abnormality presentation (6.)
    5. Point of interest reveal (9.)
11. To present the graph as a whole, each shape is visited, in sequence.
    1. If the point of interest is not included in the shape, only its shallow summary (8.) is given.
    2. If the point of interest included in the shape, the full presentation (10.) is given.

Algorithm:

Given:

Graph G of points p1…pq

Point of interest pi = (xi, yi)

PLA of G, Sg, with segments s1…sn, where each sj = pj,1…pj,w

The set of shapes in Sg, Hg, with shapes h1…hk, where each hj = sj,1…sj,w

The set of critical points Pk = pk,1…pk,m for each shape hk