

**JoVE: Science Education**  
**The Precision of Visual Working Memory with Delayed Estimation**  
--Manuscript Draft--

<b>Manuscript Number:</b>	10020
<b>Full Title:</b>	The Precision of Visual Working Memory with Delayed Estimation
<b>Article Type:</b>	Manuscript
<b>Section/Category:</b>	Manuscript Submission
<b>Corresponding Author:</b>	Jonathan Flombaum UNITED STATES
<b>Corresponding Author Secondary Information:</b>	
<b>Corresponding Author's Institution:</b>	
<b>Corresponding Author's Secondary Institution:</b>	
<b>First Author:</b>	Jonathan Flombaum
<b>First Author Secondary Information:</b>	
<b>Order of Authors:</b>	Jonathan Flombaum
<b>Order of Authors Secondary Information:</b>	

PI: Jonathan Flombaum

## Psychology Education Title

# The Precision of Visual Working Memory with Delayed Estimation

---

## Overview:

Human memory is limited. Throughout most of its history, experimental psychology has focused on investigating the discrete, quantitative limits of memory — how many individual pieces of information a person can remember. Recently, experimental psychologists have also become interested in more qualitative limits — how *precisely* is information stored?

The concept of memory precision can be at once both intuitive and elusive. It is intuitive, for example, to think that a person will remember precisely how their mother sounds, making it possible to recognize one's mother immediately over the phone or in a crowd. But how can one quantify the precision of such a memory, exactly how similar it is to the voice itself?

To study the precision of memory, and working memory in particular, experimental psychologists have devised a paradigm known as delayed estimation. It has been used most often, thus far, to study the precision of visual memories, especially memory for color. This video demonstrates standard procedures for investigating the precision of color working memory using delayed estimation.

## Procedure

### 1. Stimulus design

- 1.1. The key to any color working memory experiment is to choose a large set of individual colors that can serve as stimuli from trial to trial. Importantly, the colors need to have the same luminance (intensity of light) and the same contrast relative to the background, in order to prevent any one color from being naturally more memorable than any other. Most experiments include 180 individual colors, each with the same luminance, but varying in hue (Figure 1).

### 2. Procedure

- 2.1. Each experimental trial includes three parts (**Figure 2**):
  - A. The sample phase: One of the 180 colors is selected randomly, and presented in the display within a small square for 100 ms.
  - B. The delay: The sample disappears, and the participant faces a blank display for 900 ms.
  - C. The test: An empty square appears, along with the full color ring. The participant's task is to recall the color seen during the sample phase (part A), and to click (with the mouse) that color on the ring.

### 3. Analysis

**3.1** To understand the analysis of delayed estimation data, consider a hypothetical scenario in which the same color is the sample on many trials (Figure 3). The participant will not always respond with exactly that color. They also won't respond with ones that are very different, so the expectation is that the exactly correct color is selected most frequently. Colors very similar to it should be selected less often, but still frequently. And very different colors should be selected almost never. This kind of pattern can be described mathematically in terms of a normal distribution — a bell curve. The correct answer should be the average response over many trials, but owing to imprecision in memory, there should also be some spread. Quantifying the spread amounts to quantifying the precision of color working memory. Here is how it is done (figure 3).

**3.2** The raw data on each trial are a response color and the true target color. That means that the accuracy of a response on each trial can be quantified in terms of the angular difference between the right answer and the given answer. Recall that the colors—including the target and any response— make up a ring, occupying a total of 360 degrees. For each trial, compute the angular response error. When the answer given is exactly right, the angular error will be zero, and the most it can ever be is 358 degrees.

**3.3** Because the colors all have equal luminance and contrast, responses can be collapsed across color targets —analyzed just like they would be in the hypothetical case of a single color repeated many times. In other words, angular error amounts to a color-agnostic way of characterizing the accuracy of each response, and angular errors can be averaged across all trials under the assumption that there is no reason for any one color to produce average angular errors larger than any other.

**3.4** In this way, the responses across all trials can be used to generate a distribution of the frequency of different angular errors. Such a distribution is shown in **Figure 4**. Note that the mean of the distribution is zero and that it is normally distributed.

**3.5** From the distribution of angular errors, the precision of color working memory can be computed. Specifically, the standard deviation of the distribution is the most often used measure to describe the spread of responses. A large standard deviation means that the distribution has a lot of spread and variability, a reflection of relatively imprecise responses. A small standard deviation reflects a tight distribution, and precise memory. In this way large numbers reflect imprecision and small numbers reflect precision, so scientists often use the inverse of the standard deviation (one divided by the standard deviation) to quantify precision. Now large numbers designate precise memories and small numbers designate imprecise memories.

## Representative Result

A common reason to investigate the precision of color working memory with delayed estimation is in order to understand the relationship between discrete and continuous aspects of memory. Does remembering more individual objects produce less precise memories? The answer seems to be yes. A typical experiment will involve one or more color squares on each trial, with the participant instructed to remember all, but only tested one after the delay (**Figure 5**).

When this experiment is conducted, researchers typically compute memory precision just as described, but independently for memory load, which refers to the number of color squares in a trial.

With higher memory loads, precision tends to decline (**Figure 6**), suggesting a tradeoff between how many things a person can store in memory and how precisely they can store those things.

## Applications

Delayed estimation is a relatively new paradigm in experimental psychology, though it has become rapidly influential. In addition to investigating tradeoffs between memory capacity and precision, it can be used to compare the precision of memory systems, such as color working memory compared to color long term memory, and also to compare precision across individuals. For example, do interior decorators or painters tend to have more precise memory of color, than say lawyers or doctors?

### Legend:

**Figure 1** – A color ring including 180 individual colors. The ring is shown rendered in CIELAB space. All samples have the same  $L^*$  coordinate value, roughly meaning that they have the same luminance. The center point of the ring (shown accurately in grey) is an achromatic point, with the same luminance as the sample colors, but not chromatic value (i.e. with  $a^*$  and  $b^*$  coordinates equal to zero). The 180 individual color samples vary in terms of  $a^*$  and  $b^*$  values, specifying their proportional mixtures of blue/yellow and magenta/green to produce each individual color.

**Figure 2** – Delayed estimation procedure. In each trial, one of the 180 individual colors ('the sample') is shown for 100 ms, the display becomes blank for 900 ms, and then the participant must report the remembered sample color via mouse click on the color ring.

**Figure 3 – Hypothetical analysis assuming the same color feature on each trial.** Supposing the blue labeled 'actual target color' were the memory target over many trials, one would expect that color to obtain the most responses (i.e. the highest response frequency), with roughly a normal distribution for nearby blues. In other words, one should expect normally distributed responses with the actual target as the mean of the distribution. The variability in the distribution can supply a measure of memory precision.

**Figure 4** — Frequency of angular errors, collapsed across all trials, over the course of an experiment. Errors should form a normal distribution, centered on zero —indicating the correct response as the average answer. The variability of the distribution, specifically, the standard

deviation can be used to estimate memory precision.

**Figure 5** – Delayed estimation with four sample items. The participant is instructed to remember all four colors, but only one item is probed at test.

**Figure 6** – Memory precision as a function of memory load, the number of color samples to remember in a given trial. Note that the unit of precision is inverse degrees, ( $1/^\circ$ ), since the unit of angular response error and its standard deviation is degrees.

**PI:** Jonathan Flombaum

**Psychology Education Title:** The Precision of Visual Working Memory with Delayed Estimation

## Overview:

Human memory is limited. Throughout most of its history, experimental psychology has focused on investigating the discrete, quantitative limits of memory — how many individual pieces of information a person can remember. Recently, experimental psychologists have also become interested in more qualitative limits — how precisely is information stored?

The concept of memory precision can be both intuitive and elusive at once. It is intuitive, for example, to think a person can remember precisely how their mother sounds, making it possible to recognize one's mother immediately over the phone or in a crowd. But how can one quantify the precision of such a memory? ~~Exactly how similar~~ is the memory to it is ~~to~~ the voice itself?

To study the precision of memory and working memory, in particular, experimental psychologists have devised a paradigm known as delayed estimation. It has been used most often, thus far, to study the precision of visual memories, especially memory for color, and to understand how memory degrades the more one tries to remember at once. This video demonstrates standard procedures for investigating the precision of color working memory using delayed estimation, with a focus on how memory is affected as one tries to remember the colors of more objects simultaneously.

Choosing colors for a color working memory experiment is vital to the success of the experiment. It's important to choose colors that reside on the same mental color circle, so the colors all have the same luminance, in virtue of residing on the same plane, and the same contrast, in virtue of being equidistant from the background color. Physically, the color one perceives is related to a linear dimension, the wavelengths of light reflecting from a surface. But, perceptually, color space — the relationships in how colors are mentally represented — are non-linear. Even at the earliest ages, kids are taught to think about color "circles" and "rings."

In this video, each experimental trial includes three parts (**Figure 1**). Part A, the sample phase, where one to eight of the 180 colors is selected randomly and presented in the display, each within a small square for 500 ms; Part B, the delay, where the samples disappear, and the participant faces a blank display for 900 ms; Part C, the test, where an empty square appears, along with the full color ring. The participant's task is to recall the color seen during the sample phase (Part A) and to click with the mouse that color on the ring.

## Procedure:

**Comment [DR1]:** Feedback from JoVE 3/9:  
 \* Procedural detail - The stimulus design and experimental sections are extremely short and could use additional detail. There needs to be more explanation of subject-experimenter interaction, confounding issues that guide choice of color, and general timing of the experiment (how long does the subject have to make color selection).  
 • Data Analysis - This section should focus only on the concrete steps that will be scripted and represented in the video. There is too much material here that belongs more in a summary section.

## 1. Stimulus design.

1.1. Choose a large set of individual colors to serve as stimuli from trial to trial. The key to any color working memory experiment is to choose a large set of individual colors that can serve as stimuli from trial to trial.

1.1.1. Importantly, Make sure the colors need to have the same luminance (intensity of light) and the same contrast relative to the background, in order to prevent any one color from being naturally more memorable than any other. Physically the color one perceives is related to a linear dimension, the wavelengths of light reflecting from a surface. But Most perceptually color space — the relationships in how we mentally represent colors — are non-linear. Even at the earliest ages, kids are taught to think about color 'circles' and 'rings.'

Comment [JR2]: Moved to Overview.

1.1.2. When making the color choices, reference In a color working memory experiment, one wants to choose colors that reside on the same mental color circle, so that the colors all have the same luminance in virtue of residing on the same plane, and the same contrast, in virtue of being equidistant from the background color. CIELAB, which is an internationally standardized way of describing perceptual color space in three dimensions. This makes it easier to select colors with the right properties.

Comment [JR3]: Moved to Overview.

1.1.2.1.1.3. Select One wants to select colors that together form a circle together, with and one wants the background color as to be the center of that circle. Most experiments include 180 individual colors, each with the same luminance, but varying in hue (Figure 21).

## 2. Procedure.

### 3.

4.2. 2.1. Each experimental trial includes three parts (Figure 2):

Comment [JR4]: Moved to Overview.

4.1.2.1. Before beginning, instruct the participant to remember the stimuli and their colors. Program the experiment to do the following: On each trial, show one to eight of the 180 colors should be shown in the display for 500 ms.

4.1.1.2.1.1. Pick the colors on each trial randomly, and render each of the colors each in a small square, occupying about one degree of visual angle. These squares are the sample stimuli.

4.1.2.2.1.2. Make sure each trial has one to eight squares. Overall, the experiment should have 60 trials, each with 1, 2, 3, 4, 5, 6, 7, and 8 sample squares. This makes a total of 480 trials.

4.1.3.2.1.3. In addition to the 480 experimental trials, start the experiment with 10 practice trials. The first 5 should have only 1-2 sample items to get the participant acclimated.

4.1.3.1.2.1.3.1. At the start of the experiment, explain the instructions to the participant as follows: "In this experiment, we want to study how precisely people remember colors. On each trial, you will see a number of squares presented simultaneously in different colors. Your job is to try to remember the color of each square as precisely as you can. After about half a second, the squares will disappear. Keep them in mind. Then, the positions of one of the squares will be probed, and your job is to report the color of the square that was just in that position on a color wheel. We'll do 10 practice trials to get you used to everything. Feel free to ask questions, if you have any. Do your best, and if you feel uncertain, just guess."

4.2.2.2. After the sample squares disappear, present an empty display for 900 ms. This is the delay period, during which the participant needs to maintain their memory for sample items just seen, since they are no longer present in the display.

4.3.2.3. The delay period is followed by the test: start by randomly selecting one of the items from the sample display.

4.4.2.4. In the place where it was originally, draw a black square outline. This is the probe. It tells the participant which item to recall from memory.

4.5.2.5. Along with the probe, present the color ring that the colors were selected from.

4.6.2.6. Instruct the participant to click on the ring as close to the color as they can remember for the probed sample item.

4.6.1.2.6.1. In each test trial, present the ring in a different random rotation, so the participants can't associate specific parts of the space with specific colors.

4.6.2.2.6.2. Make sure to explain to the participant that if they feel unsure about the answer on any trial, they should guess. Leave the test display present until a response is made, with a trial ending whenever the participant clicks on the color ring.

4.7.2.7. Store as much data as possible about each trial. A number of things are critical for the output file in this experiment:



4.7.1.2.7.1. Create the output file ~~on~~will be a spreadsheet. Each line in the sheet ~~will~~ reflects a given trial.

4.7.2.2.7.2. Critically, record the following: the number of sample items in the trial, the true color of the probed item, and the color the participant selected as a response. With these data, the angular difference between the true and responded color can be computed later.

A. The sample phase: One to eight of the 180 colors is selected randomly, and presented in the display, each within a small square for 100-500 ms.

B. The delay: The samples disappear, and the participant faces a blank display for 900 ms.

C. The test: An empty square appears, along with the full color ring. The participant's task is to recall the color seen during the sample phase (part A), and to click (with the mouse) that color on the ring.

2.6. In addition to the 480 experimental trials, start the experiment with 10 practice trials. The first five should have only one or two sample items to get the participant acclimated.

2.6.1. At the start of the experiment explain the instructions to the participant as follows. "In this experiment, we want to study how precisely people remember colors. On each trial, you will see a number of squares presented simultaneously in different colors. Your job is to try to remember the color each square as precisely as you can. After about half a second, the squares will disappear. Keep them in mind. Then, the positions of one of the squares will be probed, and your job is to report the color of the square that was just in that position on a color wheel. We'll do 10 practice trials to get you used to everything. Feel free to ask questions if you have any. Do your best, and if you feel uncertain just guess."

Comment [JR5]: Moved to Overview.

Comment [JR6]: Moved to Step 2.1.3.1.

### 5.3. Analysis.

5.1.3.1. For each trial, compute the angular response error.

5.1.1.3.1.1. Compute the number of colors between the correct response in each trial and the one given, and then multiply that number by 2, since 2 degrees separate each of the colors. The result is the angular error in each trial.

5.1.2.3.1.2. Make this a column in the spreadsheet.

5.2.3.2. Average together the angular errors across all trials (under the assumption that there is no reason for any one color to produce average angular errors larger than any other). This results in a distribution of the

frequency of different angular errors (**Figure 3**). Note that the mean of the distribution is zero, and it is normally distributed.

5.3.3.3. From the distribution of angular errors, compute the precision of color working memory can be computed. Specifically, the standard deviation of the distribution is the most often used measure to describe the spread of responses. A large standard deviation means that the distribution has a lot of spread and variability, a reflection of relatively imprecise responses. A small standard deviation reflects a tight distribution, and precise memory. In this way large numbers reflect imprecision and small numbers reflect precision, so scientists often use the inverse of the standard deviation (one divided by the standard deviation) to quantify precision. Now large numbers designate precise memories and small numbers designate imprecise memories.

Comment [JR7]: Moved to Results.

3.1 To understand the analysis of delayed estimation data, consider a hypothetical scenario in which the same color is the sample on many trials (Figure 3). The participant will not always respond with exactly that color. They also won't respond with ones that are very different, so the expectation is that the exactly correct color is selected most frequently. Colors very similar to it should be selected less often, but still frequently. And very different colors should be selected almost never. This kind of pattern can be described mathematically in terms of a normal distribution—a bell curve. The correct answer should be the average response over many trials, but owing to imprecision in memory, there should also be some spread. Quantifying the spread amounts to quantifying the precision of color working memory. Here is how it is done (figure 3).

Comment [JR8]: Moved to Results.

3.2 The raw data on each trial are a response color and the true target color. That means that the accuracy of a response on each trial can be quantified in terms of the angular difference between the right answer and the given answer. Recall that the colors—including the target and any response—make up a ring, occupying a total of 360 degrees. When the answer given is exactly right, the angular error will be zero, and the most it can ever be is 358 degrees. For each trial, compute the angular response error. Compute the number of colors between the correct response in each trial and the one given, and then multiply that number by 2, since 2 degrees separate each of the colors. The result is the angular error in each trial. Make this a column in the spreadsheet.

Comment [JR9]: Moved to Results.

Comment [JR10]: Moved to Step 3.1.

3.3 2 Because the colors all have equal luminance and contrast, responses can be collapsed across color targets—analyzed just like they would be in the hypothetical case of a single color repeated many times (Figure 3). The participant will not always respond with exactly that color. They also won't respond with ones that are very different, so the expectation is that the exactly correct color is selected most frequently. Colors very similar to it should be selected less often, but still frequently. And very different colors should be selected almost never. This kind of pattern can

be described mathematically in terms of a normal distribution — a bell curve. The correct answer should be the average response over many trials, but owing to imprecision in memory, there should also be some spread. Quantifying the spread amounts to quantifying the precision of color working memory (**figure 3**). Because the colors all have equal luminance and contrast, responses can be collapsed across color targets — analyzed just like they would be in the hypothetical case of a single color repeated many times.

In other words, angular error amounts to a color-agnostic way of characterizing the accuracy of each response, and angular errors can be so averaged together the angular errors across all trials (under the assumption that there is no reason for any one color to produce average angular errors larger than any other).

**3.4** In this way, the responses across all trials can be used to generate a distribution of the frequency of different angular errors. Such a distribution is shown in **Figure 4**. Note that the mean of the distribution is zero and that it is normally distributed.

**Comment [JR11]:** Moved to Results.

**Comment [JR12]:** Moved to 3.1.2 and 3.2.

## Representative Results:

The raw data on each trial are a response color and the true target color. That means the accuracy of a response on each trial can be quantified in terms of the angular difference between the right answer and the given answer. The colors — including the target and any response — make up a ring, occupying a total of 360 degrees. When the answer given is exactly right, the angular error is zero, and the most it can ever be is 358 degrees.

Because the colors all have equal luminance and contrast, responses can be collapsed across color targets — analyzed just like they would be in the hypothetical case of a single color repeated many times (**Figure 4**). The participant does not always respond with exactly that color. They also won't respond with ones that are very different, so the expectation is that the exactly correct color is selected most frequently. Colors very similar to it should be selected less often, but still frequently. And very different colors should be selected almost never. This kind of pattern can be described mathematically in terms of a normal distribution — a bell curve. The correct answer should be the average response over many trials, but owing to imprecision in memory, there should also be some spread — quantifying the spread amounts to quantifying the precision of color working memory (**Figure 4**).

In other words, angular error amounts to a color-agnostic way of characterizing the accuracy of each response.

The standard deviation of the distribution is the most often used measure to describe the spread of responses. A large standard deviation means that the distribution has a lot of spread and variability, a reflection of relatively imprecise responses. A small standard deviation reflects a tight distribution and precise memory. In this way, large numbers

reflect imprecision and small numbers reflect precision, so scientists often use the inverse of the standard deviation (one divided by the standard deviation) to quantify precision. Now, large numbers designate precise memories, and small numbers designate imprecise memories.

A common reason to investigate the precision of color working memory with delayed estimation is in order to understand the relationship between discrete and continuous aspects of memory. Does remembering more individual objects produce less precise memories? The answer seems to be yes. A typical experiment will involve one or more color squares on each trial, with the participant instructed to remember all, but only tested one after the delay (Figure 5).

When this experiment is conducted, researchers typically compute memory precision just as described, but independently for memory load, which refers to the number of color squares in a trial.

With higher memory loads, precision tends to decline (Figure 56), suggesting a tradeoff between how many things a person can store in memory and how precisely they can store those things.

### Applications:

Delayed estimation is a relatively new paradigm in experimental psychology, though it has become rapidly influential. In addition to investigating tradeoffs between memory capacity and precision, it can be used to compare the precision of memory systems, such as color working memory compared to color long term memory, and also to compare precision across individuals. For example, do interior decorators or painters tend to have more precise memory of color than lawyers or doctors?

### Legend:

Figure 1: Delayed estimation procedure. In each trial, one of the 180 individual colors (the sample) is shown for 100 ms, the display becomes blank for 900 ms, and then the participant must report the remembered sample color via mouse click on the color ring.

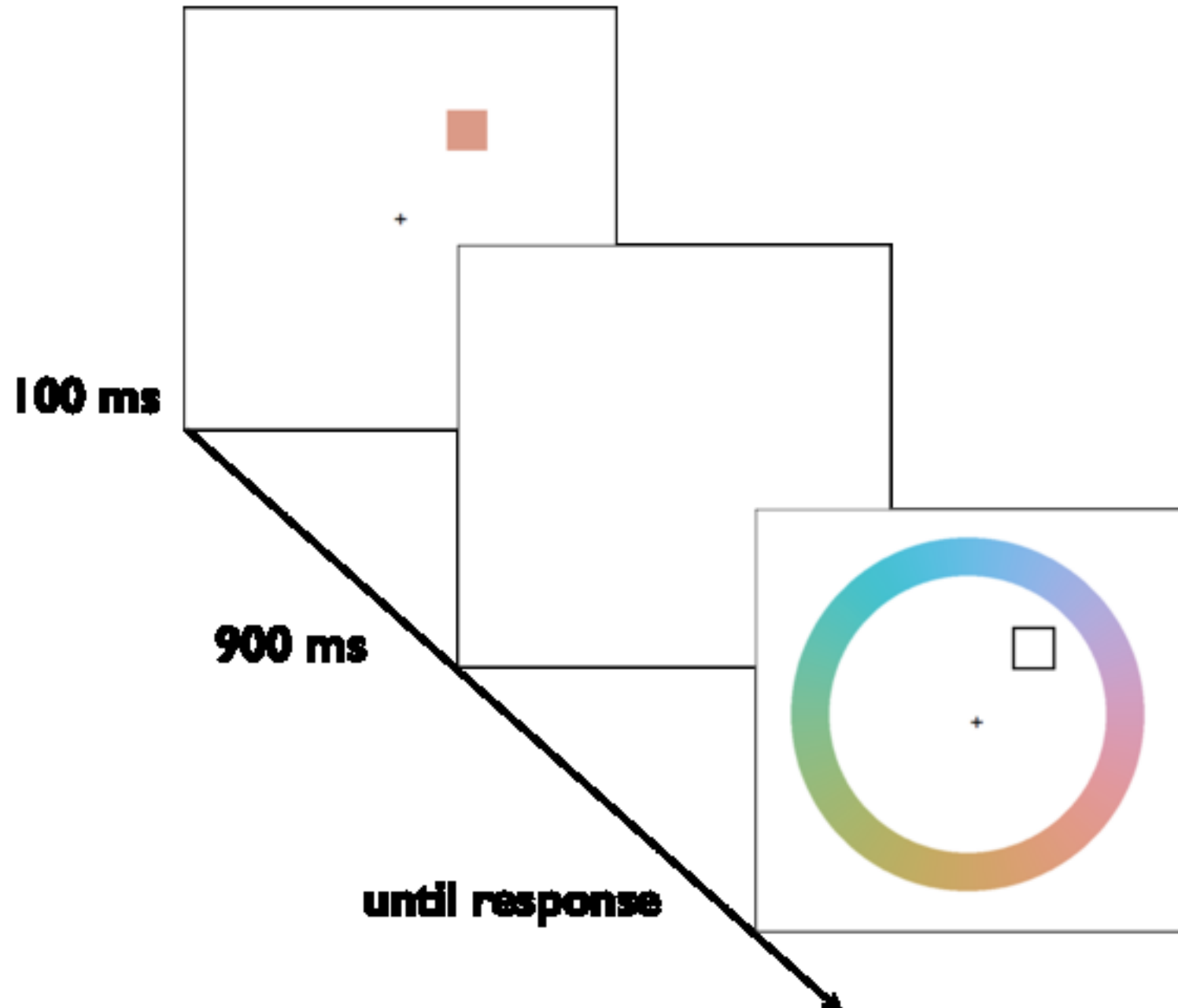
Figure 2: A color ring including 180 individual colors. The ring is shown rendered in CIELAB space. All samples have the same  $L^*$  coordinate value, roughly meaning that they have the same luminance. The center point of the ring (shown accurately in grey) is an achromatic point, with the same luminance as the sample colors, but not chromatic value (i.e., with  $a^*$  and  $b^*$  coordinates equal to zero). The 180 individual color samples vary in terms of  $a^*$  and  $b^*$  values, specifying their proportional mixtures of blue/yellow and magenta/green to produce each individual color.

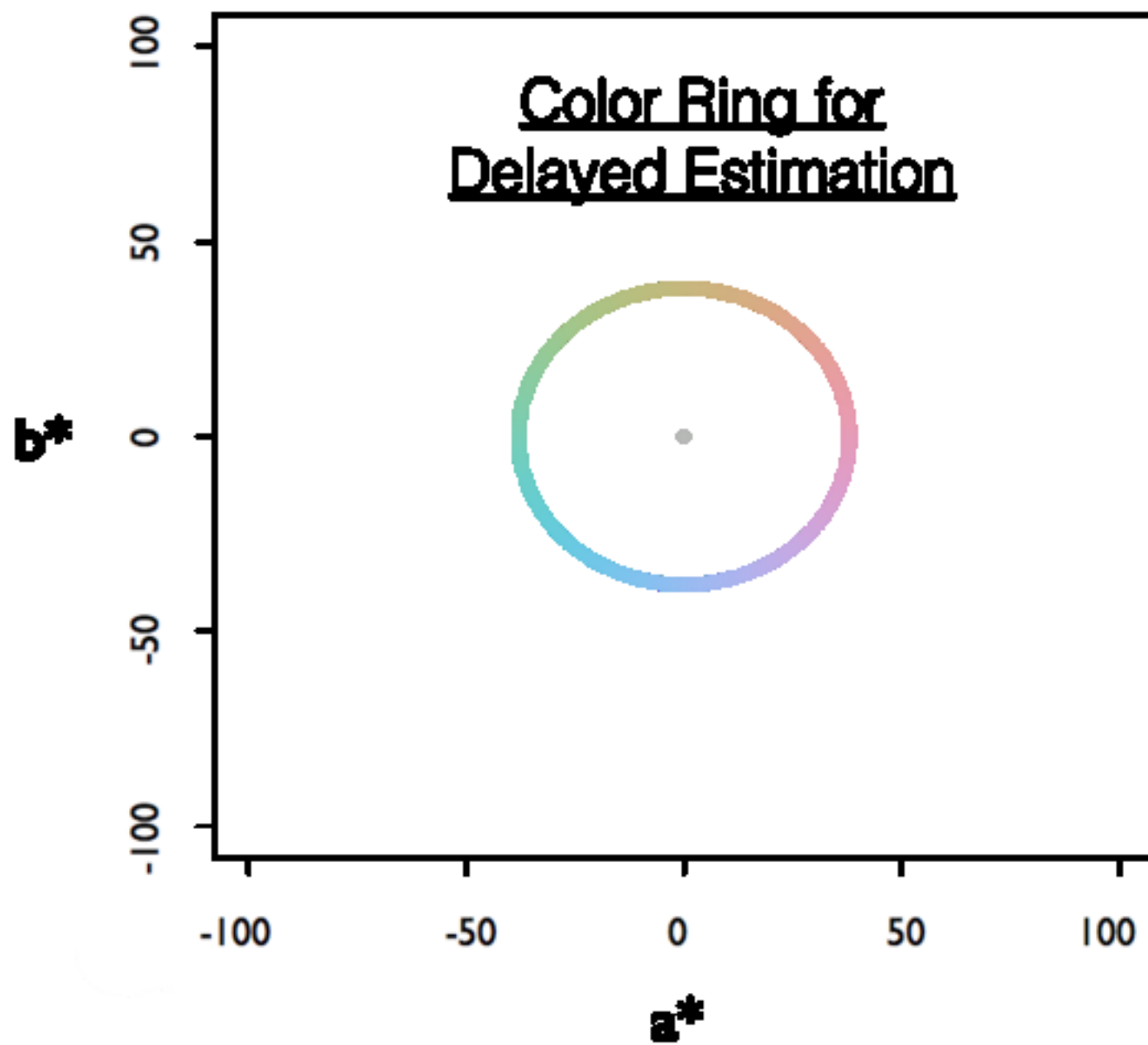
Figure 3: Frequency of angular errors, collapsed across all trials, over the course of an experiment. Errors should form a normal distribution, centered on zero — indicating the correct response as the average answer. The variability of the distribution, specifically, the standard deviation can be used to estimate memory precision.

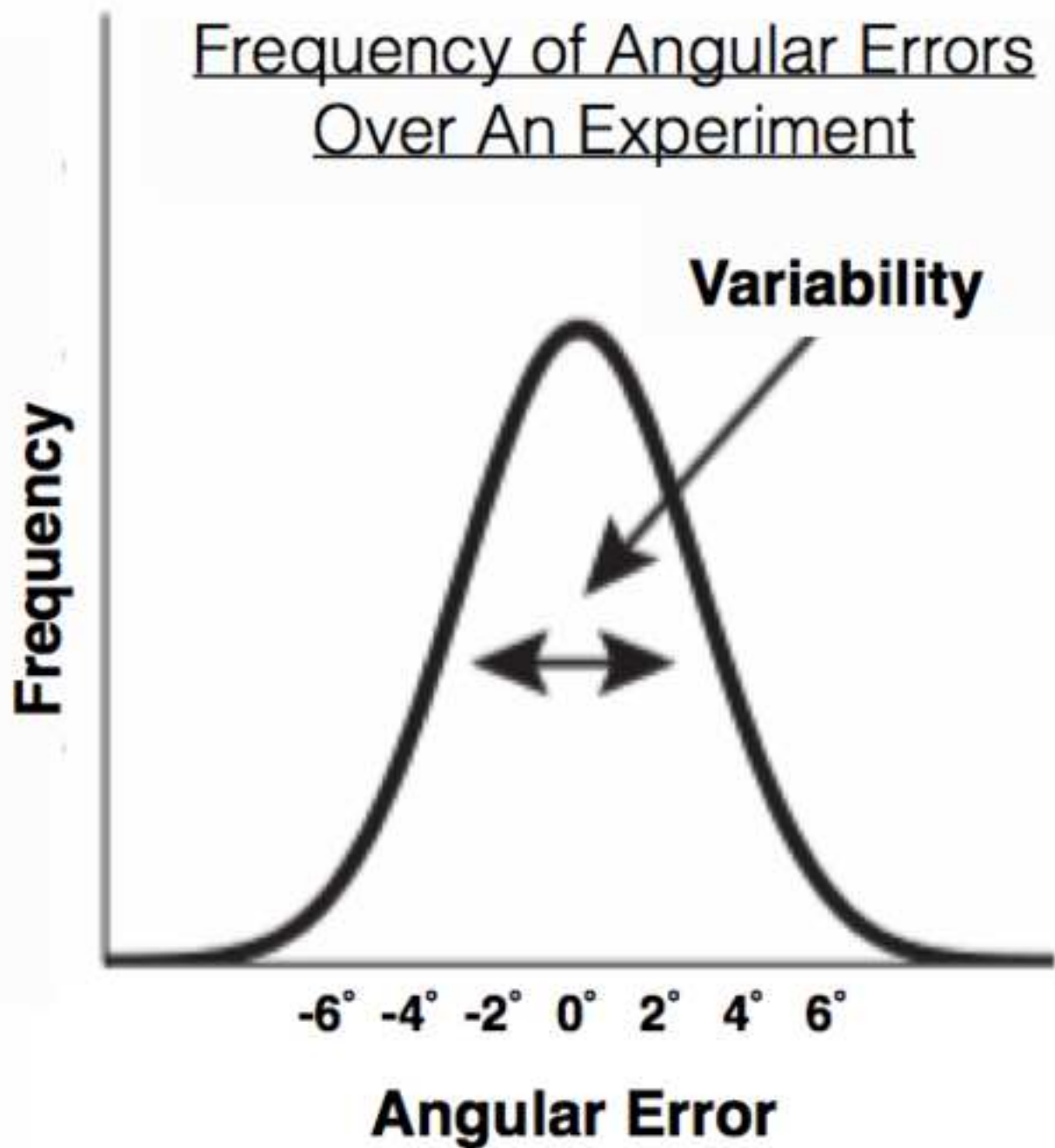
Figure 4: Hypothetical analysis assuming the same color feature on each trial. Supposing the blue labeled “actual target color” were the memory target over many trials, one would expect that color to obtain the most responses (i.e., the highest response frequency), with roughly a normal distribution for nearby blues. In other words, one should expect normally distributed responses with the actual target as the mean of the distribution. The variability in the distribution can supply a measure of memory precision.

Figure 5: Memory precision as a function of memory load, the number of color samples to remember in a given trial. Note that the unit of precision is inverse degrees, ( $1/^\circ$ ), since the unit of angular response error and its standard deviation is degrees.

# Delayed Estimation Procedure



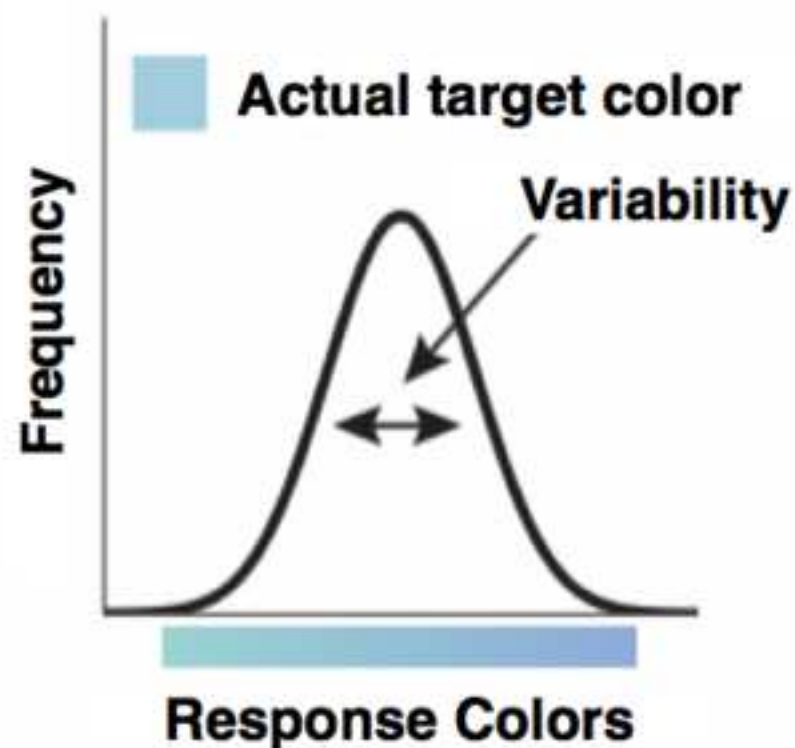






## Hypothetical Analysis Assuming the Same Color Target on Each Trial

**Six individual responses  
to the same target**



# Memory Precision as a Function of Memory Load

