Utrecht University
Bachelor Thesis

The Leaning Tower Illusion

A conflict between two- and three-dimensional parallelism

Author:
Carel van Leeuwen (3510336)

Supervisor:
Prof. Dr. Casper Erkelens

25 June 2014

Abstract:

The Leaning Tower Illusion is a visual illusion in which two identical images, originally of the leaning tower of Pisa, appear to lean away from each other when juxtaposed. The illusion is explained as perspective cues distorting our sense of 2-dimensional direction.[1] We used human subjects to test this explanation and to explore the way in which lines create a sense of perspective. We predict that the strength of the illusion scales linearly with the horizontal distance between the image, and is independent of their vertical length. The results turn out to be inconclusive because of large standard deviations in individual measurements. We do show how different compositions of lines, or even a single line, are seen in perspective and give rise to the notion of depth.

Table of Contents

1. Introduction

The Leaning Tower Illusion is a visual illusion that occurs when two identical images containing perspective cues are placed next to each other. The illusion was discovered by Frederick Kingdom, Ali Yoonessi and Elena Gheorghiu at McGill University, Canada and was awarded first prize in the Best Visual Illusion of the Year contest in 2007.

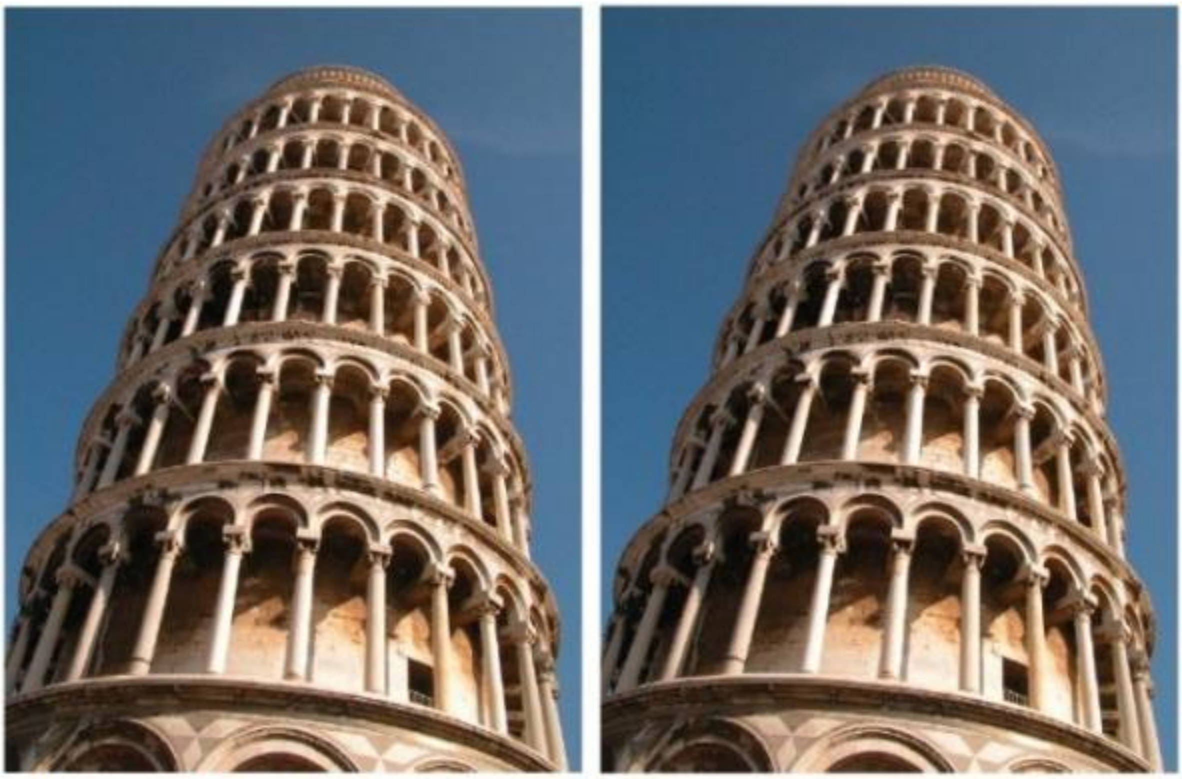


Figure 1: The original Leaning Tower Illusion.

The two images of the leaning tower of Pisa are identical, yet the towers appear to lean away from each other. The explanation is straightforward: If two lines are parallel in the three-dimensional world and one looks along the direction of the lines, one sees the lines converging in a single vanishing point. In this illusion an image is simply copied and moved, and the vanishing point of lines in the image moves along with it. Lines disappearing in different vanishing points in our retinal image represent lines in 3D space that are not parallel but diverge, hence the towers appear to diverge.

2. Research Questions and Hypotheses

While this perspective explanation seems quite simple, we wanted to test whether it is actually correct. To do this, we asked ourselves a number of questions which could be tested by experiment and for which the perspective explanation predicts an answer.

- How does the effect depend on the horizontal distance between the images?
- How does the effect depend on the vertical length of the images?
- How does the effect depend on the perspective cues within the image?

The answer to the first question is straightforward: The different vanishing points in the 2D image correspond to diverging directions in 3D space. The bigger the distance between the vanishing points, the more divergent the 3D directions appear. Therefore the perceived leaning should scale linearly with the horizontal distance between the images.

As for the second question, two views make a compelling argument: One could say that a longer vertical length of the image decreases the perceived depth, because the depth gradient decreases. This is similar to the way that the perceived speed of the car you're driving decreases if the car is tall and your eyes are positioned high above the road. However the horizontal distance between the vanishing points of the objects in the image does not change as the image gets taller. For this reason we do not expect the strength of the illusion to depend on vertical image length.

The third question is easy to answer in theory but difficult to test. Obviously the whole effect depends on perspective cues; if one takes an image without perspective cues, copies it and puts it next to the original, the effect does not occur. For this reason we expect the effect to increase when the image contains more perspective cues. This is however difficult to test experimentally because what might appear as a perspective cue to some viewer, may appear to another as just a bunch of lines. Even a single viewer can change his or her opinion on this once the perspective has been pointed out to him or her. We do expect the effect to be non-existent without perspective cues.

Figure 2: Without perspective cues, no leaning is perceived.[2]

3. Research Methods

A total of 9 subjects were used. All had normal or corrected-to-normal vision. Subjects were aged between 20 and 25 years and were students at Utrecht University, all from the physics, mathematics and computerscience departments. They were not familiar with of the Leaning Tower Illusion. They were seated in front of a computer screen (LaCie 321 LCD monitor) at a distance of 57 cm. They placed their head upon a chinrest slightly to the right of the middle of the screen. Horizontally, the chinrest had a fixed position. Vertically they could adjust the height of the chinrest to their comfort.

Subjects were given a computer mouse and were shown a stimulus containing two straight red lines. The right red line was fixed. The left red line had one fixed end. The other end of the left red line could be moved horizontally by moving the mouse pointer. Subjects were instructed to position the left red line such that it was parallel with the right one, and click the left mouse button at that point. Upon clicking the location of the mouse was registered and subjects were presented with a pausing image for 4 seconds, after which next stimulus was shown. The pausing image was a part of a Donald Duck poster with many details but hardly any perspective cues. Subjects were not given any feedback on their performance. The stimuli and clicking environment were created with help of Wolfram Mathematica 9.0

4. Stimuli

Figure 3 shows the stimuli we used. They were created in order to isolate effects of the horizontal distance between images, the vertical length of the image and the amount of perspective cues within the image.

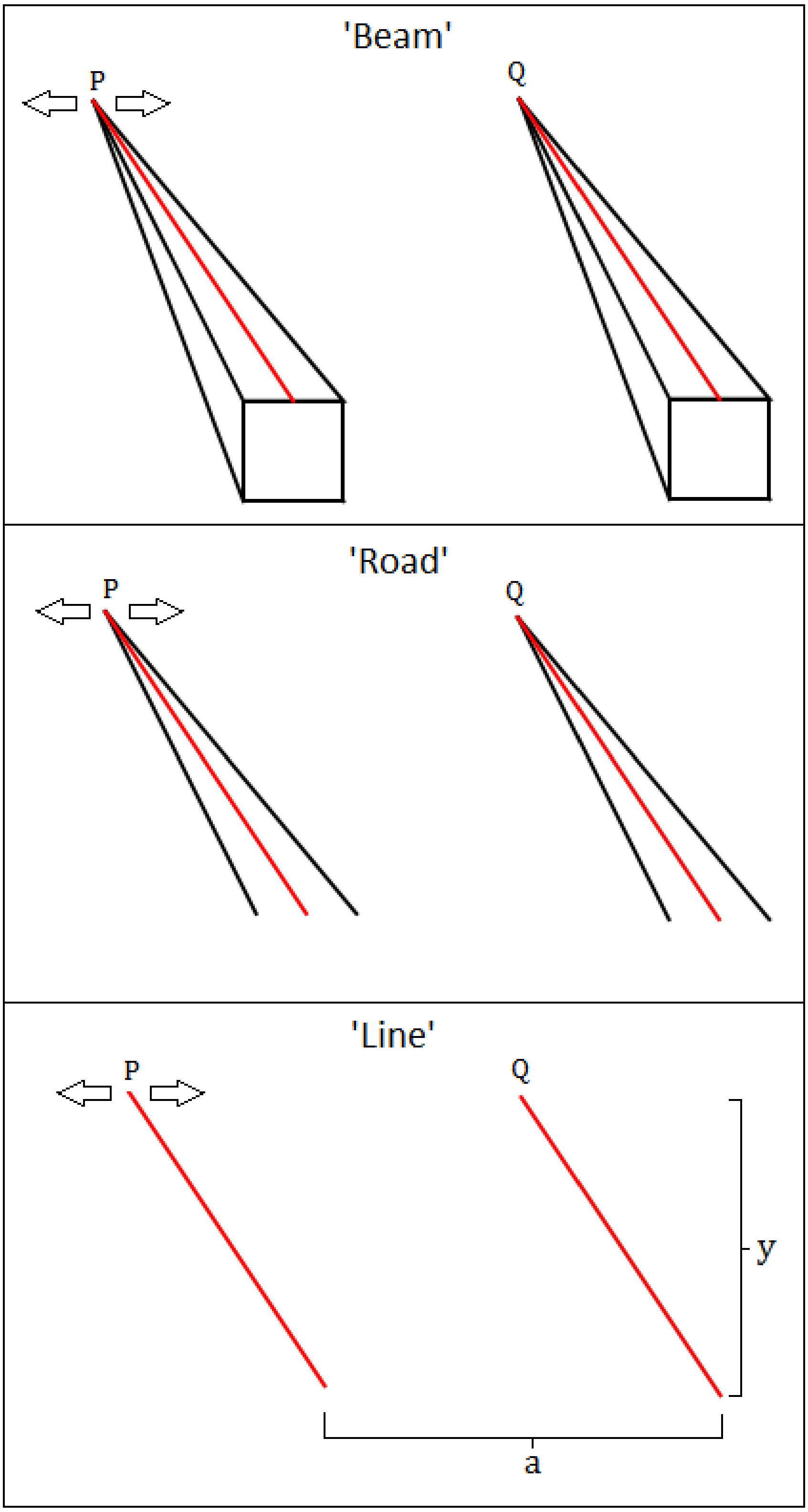


Figure 3: Three different kinds of stimuli were used. they will be referred to as 'Beam', 'Road' and 'Line'.

The right part was always fixed, with vanishing point Q straight ahead of the testsubject.

The left part could be changed by moving point P along the horizontal axis while the bottom part was fixed.

The Beam stimulus was designed to induce a stronger sense of depth than the Road stimulus. The Road represents a flat object in 3D space. The Beam consists of more lines, and represents a 3D object with a volume. The Line stimulus was designed to contain no perspective cues, and serves as a control stimulus.

We varied two parameters: The horizontal distance between the two parts of the stimulus '*a*', and the vertical length of the stimulus '*y*'.

We used 6 different values of *a* and 3 values of *y*. With the 3 types of stimuli subjects were presented a total of 54 (= 3 x 6 x 3) different images.

In Figure 3 the red lines are parallel. However, the perspective cues in the Beam and Road image distort what we perceive as parallel. If the Beams and Roads are seen as 3-dimensional shapes, the different vanishing points indicate that in 3D the red lines are divergent. In order to make the lines parallel in 3D space, one would have to move point P all the way to the right until it coincides with point Q. For this reason subjects were not positioned in front of the middle of the screen, but with point Q straight ahead of them.

The difference between where point P was put, and where it would be if the red lines were parallel, was called *d*. The ratio of the distance between the bottom ends of the red lines (*a*) and the top ends (|**P – Q**| = *a* – *d*) serves as a quantitative measure of how perspective affects depth perception.

In the experiment the first stimulus shown was always a Line-stimulus. After that the different types and parameters were randomly ordered.

We varied parameter *a* from a_1 = 10.8 cm to a_6 = 24.3 cm with regular intervals. Parameter *y* took three different values: y_1 = 21.6 cm, y_2 = 25.65 cm, y_3 = 29.7 cm. To give an idea of what this looks like on screen, figure 4 shows two Beam stimuli in a single figure. One stimulus with both parameters at minimal value (a_1,y_1), the other stimulus with both parameters at maximum value (a_6,y_3). The images in figure 4 are scaled versions of the stimuli presented in the experiment.

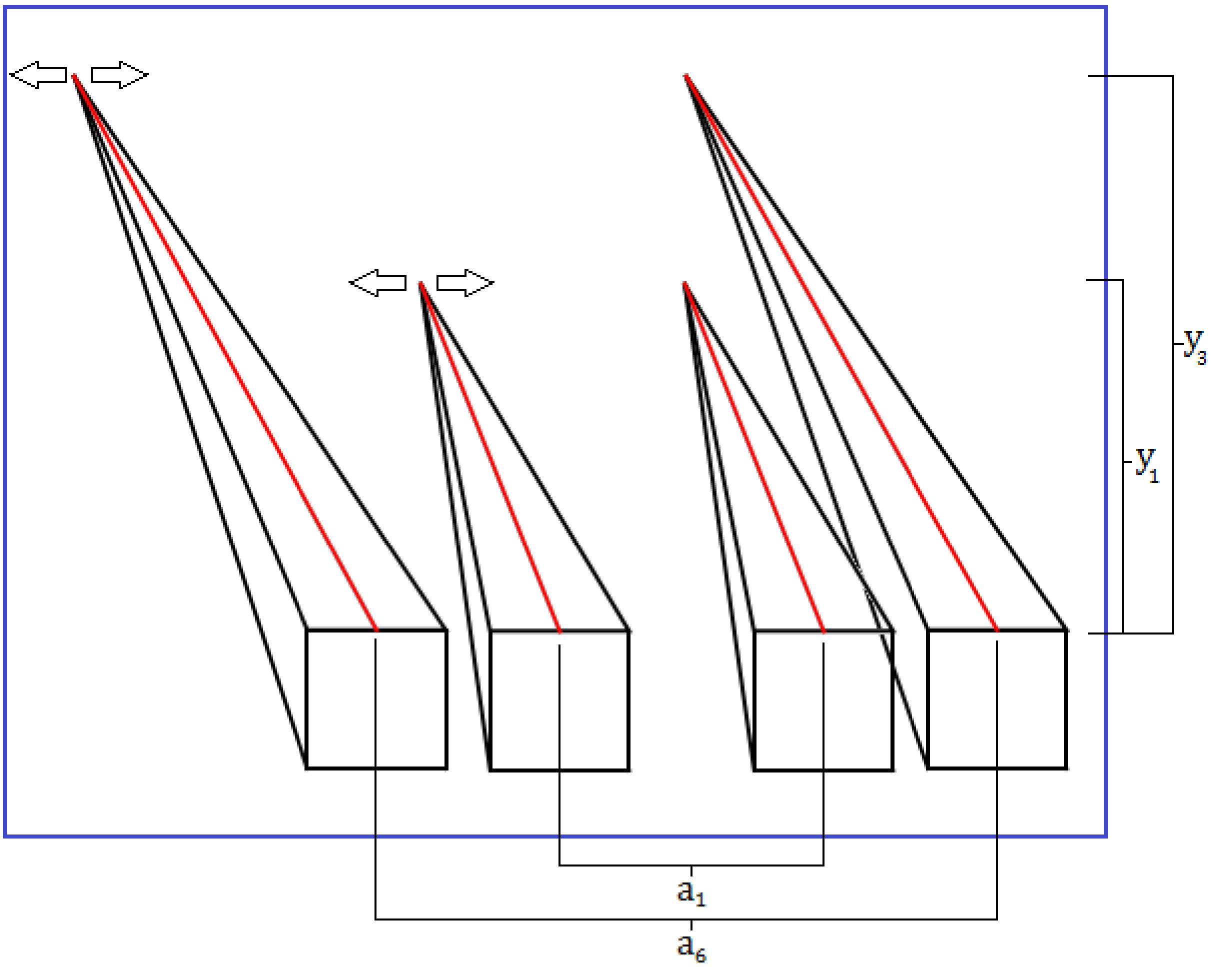


Figure 4: Two Block stimuli are shown in one image with different parameters a and y. The blue box indicates the edge of the screen. The arrows here indicate that the left vanishing point was moveable, the actual stimulus contained no arrows.

5. Results

We expect variable *d* to depend linearly on parameter *a* and to be independent of parameter *y*. This means that we expect the values of *d/a* to be normally distributed about a mean value. Figure 5 shows the distribution of all data for the three different types of stimuli, compared to a normal distribution.

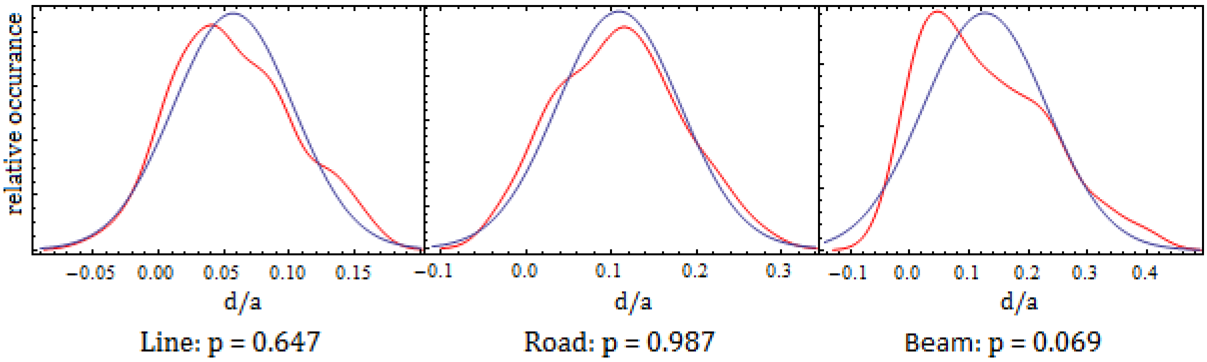


Figure 5: The red line is a probability density function of *d/a* for all different values of *a* and *y*. The blue line is a normal distribution with the same mean and standarddeviation. the p-values reflect the probability that the data came from a normal distribution and were calculated using the Anderson-Darling test.

As you can see, the results from Line- and Road-stimuli appear to be normally distributed. The Beam-stimuli results do not resemble a normal distribution. Possible reasons for this will be discussed under the Discussion section.

To test whether *d* depends linearly on *a* we used a least-squares method to construct a linear model that fits the data of *d/a* for different values of *a*. The right-hand side of figure 6 shows the data for all values of *y* and the best fitting line through the data. The left-hand side shows the slopes of the best fitting lines for different values of *y*. We expected the slopes to be zero. Assuming the slope followed a normal distribution (with $\mu = 0$ and σ = the standard deviation of the slope of the fitted line), we calculated the probability to find the discrepancies we found. For the Line-stimuli, the slope differed significantly from zero ($p = 0.021$). For the Road-stimuli, the slope was not significantly different from zero ($p = 0.762$). For the Beam-stimuli, this probability was not calculated because the Beam data was not normally distributed. However, the measurements had large standard deviations (relative SD's ranged from 0.6 and 1), but the fitted values lay very close to the mean values. For all fits the value of the reduced χ^2 was much lower than 1 (ranging from 0.02 to 0.001). Because of this, the data is inconclusive as for whether *d* scales linearly with *a*.

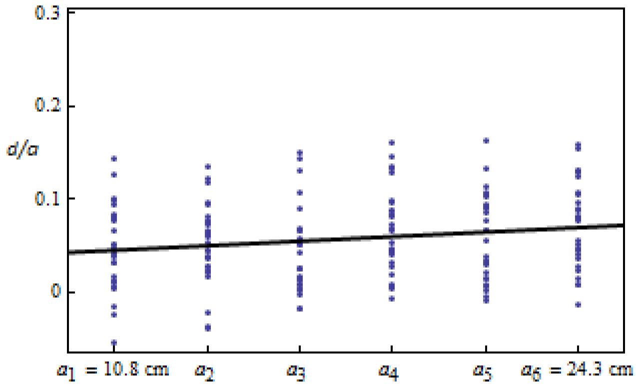
To test whether *d* is independent of *y* we used the same method as outlined above. Results are shown in figure 7. For the Line-stimuli the slope was significantly different from zero ($p = 0.0001$). For the Road-stimuli the slope was not significantly different from zero ($p = 0.229$). Again the reduced χ^2 had values below 0.02.

For all fitted lines we examined whether the studentized fit residuals were normally distributed using the Anderson-Darling test. For the Line- and Road stimuli this was always the case (all p-values > 0.7). For the Beam-stimuli this was not the case (all p-values < 0.1). On average, the Road-stimuli did return higher values of *d* than the Line-stimuli. We used a T-test to compare all values of *d/a* for the two types of stimuli. They turned out to be significantly different ($p \sim 10^{-12}$).

Line:

	Slope [cm^{-1}]	Standard Deviation
y_1	0.00162	0.00123
y_2	0.00033	0.0013
y_3	0.00345	0.00133
all y's	0.0018	0.00078

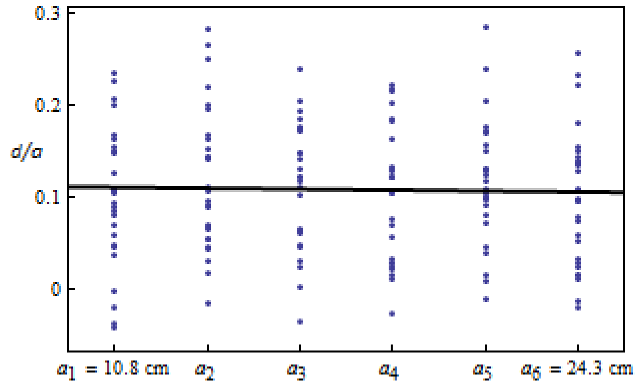
p = 0.021



Road:

	Slope [cm^{-1}]	Standard Deviation
y_1	0.00129	0.00217
y_2	−0.00004	0.00205
y_3	−0.00237	0.00221
all y's	−0.00037	0.00123

p = 0.762



Beam:

	Slope [cm^{-1}]	Standard Deviation
y_1	−0.00181	0.00351
y_2	0.00085	0.00314
y_3	0.00298	0.00259
all y's	0.00067	0.00178

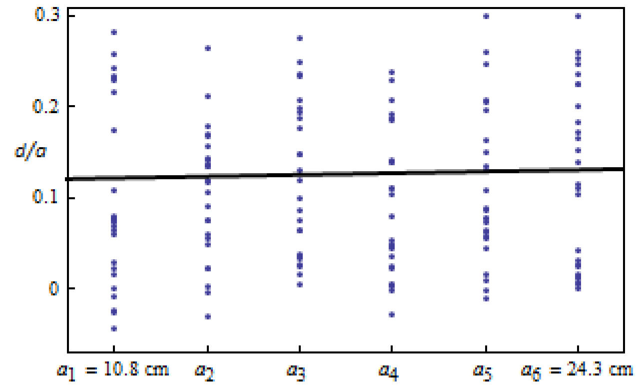


Figure 6: On the right there are plots of the data of d/a for all values of y, and the best fitting line through them. On the left there are tables of the slope for the best fitting lines, for different values of y seperately, and all together. The p-values apply to the slopes with all y's included. The p-values represent the probability to find such a non-zero slope, assuming the slope is zero and the data is normally distributed. In the case of the Beam-stimuli, this p-value is omitted because the Beam-data was not normally distributed.

Line: $\boxed{p = 0.0001}$

Slope: 0.0021 ± 0.0017 cm^{-1}

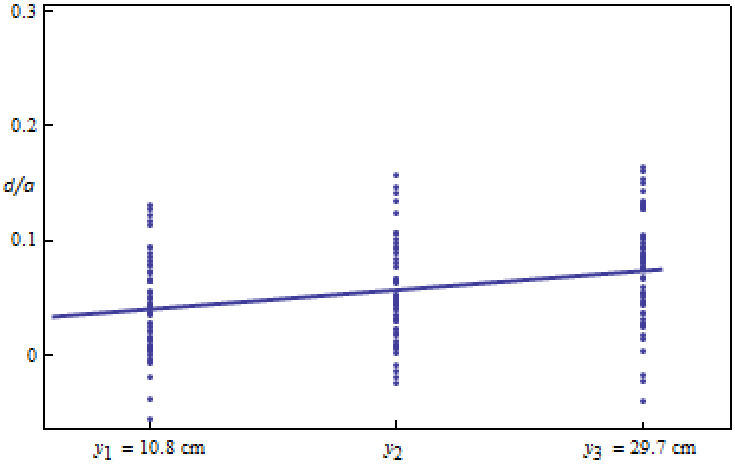
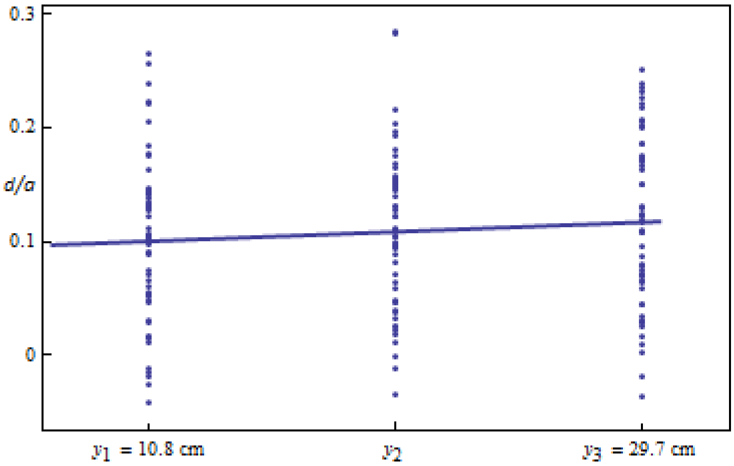


Figure 7: Here are shown the values of *d/a* at different values of *y* for all three stimuli, along with the best fitting line. The p-values reflect the probability to find the nonzero values for the slope, assuming it was zero. For the Beam-stimuli no p-value was calculated because the data was not normally distributed.

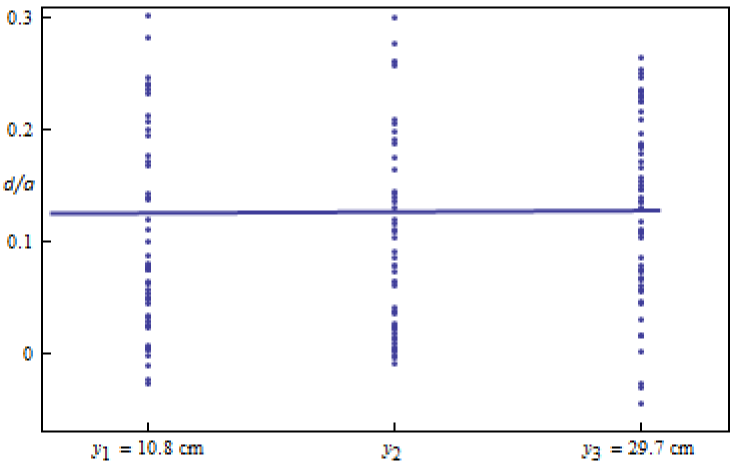
Road: $\boxed{p = 0.229}$

Slope: 0.0041 ± 0.0011 cm^{-1}



Beam:

Slope: 0.0003 ± 0.0025 cm^{-1}



6. Discussion & Conclusions

The first hypothesis, that the leaning effect should scale linearly with the horizontal distance between images, and hence $d \propto a$, does not follow from the data, because the standard deviations are too large. The second hypothesis, that *d* is independent of *y* also cannot be proven for the same reason. The low values for the reduced χ^2 indicate that it is possible that the fitted models are not reliable. We can only conclude that, in the case of the Road-stimuli, it is not impossible that the first two hypotheses are correct.

In the case of the Line-stimuli, we expected the mean value of *d* to be zero. This was not the case. In the presence of perspective cues, the top of the image is perceived as further away than the bottom. This gives rise to the perception of depth and the resulting leaning effect. However, the non-zero average of *d* from the Line-stimuli indicates that the top of the image is still seen as further away, even without perspective cues. A possible reason for this is that the height of an object in the field of vision is in itself a perspective cue. In natural images, it's logical to assume that objects which are closer to the horizon are further away. If no horizon is seen, a higher object is assumed to be further away.[3] Other cues such as linear perspective, object size or texture gradients provide much more accurate information about distance. A possible explanation is that the height in field of vision is only relevant in the absence of other distance cues. This would explain the slight increase of *d* with increasing *y* in the Line-stimuli, but not with the others.

The last hypothesis was that the perceived depth would increase if the image contained more perspective cues. We created the Beam-stimuli to test this by comparing Beams and Roads. However, the Beam-stimuli turned out to be unsuitable for this type of test. An explanation is that because of the number of lines in the Beam-stimulus, subjects were able to deduce information about whether lines were parallel, instead of just looking at the red lines. An alternative way to test the depth induced by the Beam-stimulus would be a two-alternative forced choice setup with limited time. Another option would be to rotate the square front face of one of the beams in the image. This way the left and right image are no longer exactly the same in the parallel position.

Returning to the last hypothesis, we can conclude that the converging lines of the Road-stimulus give a more powerful notion of depth than just a Line. Even a single line can induce the perception of depth. After all, the Leaning Tower Illusion could serve as a powerful tool to give a quantitative measure of how the perception of depth is affected by perspective cues.

7. References

1. Kingdom F A A, Yoonessi A, Gheorghiu E, 2007, "The Leaning Tower illusion: a new illusion of perspective" *Perception* **36**(3) 475 – 477
2. Image was made by Akiyoshi Kitaoka from Ritsumeikan University, Kyoto, to test whether the illusion could be explained by acute angle expansion.
   Source: http://www.scholarpedia.org/article/Leaning_tower_illusion
3. Ian P. Howard, 2012, "Perceiving In Depth" ISBN: 978-0-19-976416-7.
   Section 26.4, Volume 3, pages 39-40.