Falling Skyscrapers: When Cross-Modal Perception of Verticality Fails
Chia-huei Tseng, Hiu Mei Chow and Lothar Spillmann
Psychological Science published online 30 May 2013
DOI: 10.1177/0956797612472344

The online version of this article can be found at:
http://pss.sagepub.com/content/early/2013/05/23/0956797612472344

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Association for Psychological Science

Additional services and information for Psychological Science can be found at:
Email Alerts: http://pss.sagepub.com/cgi/alerts
Subscriptions: http://pss.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

>> OnlineFirst Version of Record - May 30, 2013

What is This?
Humans depend on accurate verticality perception to interact with the world. In most instances, the world appears to remain upright even when people move around, sit, or lie down. This is because the brain takes into account the observer’s body position and frame of reference in providing a representation of the visual world. In certain cases, however, observers’ sense of verticality is distorted. For example, when observers lie on the floor of a house with a frame and interiors that tilt 40° to 50° away from the gravitational vertical plane and look up, they adopt the tilted frame of reference and feel themselves sliding down the slope (Shimojo, 2008). Similarly, the famous haunted swing illusion gives stationary observers the sense of swinging around 360° when, in reality, it is the house that rotates (Metzger, 1936/2006). To elucidate the sensory origin of this misperception, we conducted field experiments on the moving tram to systematically disentangle the contributions of four sensory systems known to affect verticality perception, namely, vestibular, tactile, proprioceptive, and visual cues. Our results refute the intuitive assumption that the perceived tilt of the buildings is based on visual error signals and demonstrate instead that a unified percept of verticality is a product of the synergistic interaction among multiple sensory systems and the contextual information available in the real world.

Keywords
perception, cognitive processes

Received 8/24/12; Revision accepted 11/21/12

Humans discover the striking verticality misperception on the Hong Kong Peak Tram, a celebrated tourist attraction that has received millions of visitors from all over the world. The skyscrapers passing by the tram’s window appear to fall toward the nearby Hong Kong Victoria Peak, causing passengers to gasp at the unusual sight (Figs. 1a and 1b). One of the great differences between this illusion and the tilted room or haunted swing illusion is that the mountain tram is surrounded by open windows that allow observers to fully access the visual cues from the outside world, which prompts them to perceive the direction of gravity veridically. This observation poses the question: Why do observers abandon veridical gravity-orientation cues in favor of a misleading frame of reference?

To address this question, we ran field experiments on the moving tram and quantified the perceived visual tilt as a function of the mountain’s slope. We found that

Corresponding Author:
Chia-huei Tseng, Department of Psychology, The University of Hong Kong, Room 627, 6/F, The Jockey Club Tower, Centennial Campus, Pokfulam Road, Hong Kong
E-mail: ch_tseng@alumni.uci.edu
Fig. 1. The Hong Kong Peak Tram illusion and results from Experiment 1. The photo (a) shows how the skyscrapers outside the tram window appear to lean toward Hong Kong Victoria Peak. In the schematic diagram of the illusion (b), the angle $\theta$ illustrates the approximate deviation of the perceived tilt relative to gravity for a mountain slope of $\alpha$. The scatter plots (with best-fitting regression lines) show the perceived tilt of nearby skyscrapers for each observer as a function of the slope of the mountain, separately for nighttime and daytime. The head and back of the observers were oriented orthogonally to the tram’s slope, while their buttocks and feet were parallel to the tram’s slope.
misperception of tilt was greatest at night, and subsequent experiments provided evidence that vestibular, tactile, and proprioceptive cues also play crucial roles in forming a unified percept of verticality.

Method

Tilt-matching task

To quantify the illusory tilt, we asked 4 trained observers (22–74 years old) to sit next to the window in the Hong Kong Peak Tram during its normal operation. Observers were given a handheld chart with 13 slanted lines ranging from 0° to 36°; these lines were arranged in a different random order for each trip. Observers told a coder which line best matched their perceived tilt on each trip. Both the observers and the coder were unaware of the exact angle of particular lines. The coder also simultaneously recorded the mountain’s slope using a rotary pitch or an iPhone application (either TiltMeter or iHandy Level) attached to the tram’s window frame. When there was no illusory tilt, observers invariably chose the 0° line. Each tram ride took about 6 min, and 10 to 12 reports on average were made on each trip. For each observer, all measurements were taken during two round trips (i.e., twice while ascending and twice while descending) for each experiment. The slope of the tramline averaged 18.7° (range = 4°–27°) from the lower Garden Road terminus to the Peak Tram station.

Experiments 1 and 2: visual cues

In Experiment 1, observers’ illusory tilt was measured at both daytime and nighttime. In Experiment 2, to test how much the tram’s oblique window frame contributed to the perceived tilt, we also asked observers to place a cardboard box with a round 10-cm viewing aperture over their head, through which they viewed the buildings (Fig. 2b). Experiment 2 was conducted only at night. In both experiments, observers rested their head and back against the compartment seat, which was bolted down to the bottom floor of the tram.

Experiments 3, 4, 5, 6, and 7: nonvisual cues

We next performed a systematic study to determine how nonvisual modalities affect the Hong Kong Peak Tram illusion. These nonvisual modalities were, first, vestibular information from the otoliths in the inner ears; second, tactile information from the pressure on the back, buttocks, and feet resting on tilted surfaces; and third, proprioceptive information from the ankle joints supporting the body weight. All of these modalities produce positional signals relative to gravity due to sitting in a reclining position, which may cause observers to misperceive verticality. To discern the relative contributions of these sources, we tested each of them in isolation and in combination by using compensating wedges.

Specifically, three 18.7° Styrofoam wedges were used. One was placed behind the back (Experiment 3), one under the buttocks (Experiment 4), and one beneath the feet (Experiment 5). In Experiment 6, all three wedges were used together. The back wedge approximately restored the observer’s torso and head from the upright position to the gravitational vertical position (i.e., resulting in a change in vestibular information), thereby compensating for most of the mountain slope. The wedge placed under the buttocks normalized the pressure from the tilted seat (resulting in a change in tactile information), and the wedge inserted under the observer’s feet straightened out the ankle joints (resulting in a change in proprioceptive information). Observers performed the same tilt-matching task with the wedges applied individually and jointly (Fig. 2c). To determine whether sensory inputs differ when they are used passively (e.g., sitting down) or actively (e.g., standing up), we also conducted an additional experiment (Experiment 7), in which observers stood up while reporting the perceived tilt of the skyscrapers (Fig. 2d). Experiments 3 through 7 were all conducted at night.

Experiment 8: combined visual and nonvisual cues

In a final experiment, we studied all the visual and nonvisual cues examined in the previous experiments together by having observers wear the same box as in Experiment 2 and stand up on a foot wedge (Fig. 2e). All observers completed the experiment at night.

Results

In all experiments, the data for ascending and descending directions did not differ and therefore were collapsed in the subsequent analysis. The statistical analysis of each experiment was based on the pooled data from all 4 observers, whose individual results are plotted in the figures. The nighttime condition of Experiment 1 was used as a baseline against which all other experiments were compared. All regression lines pass origin (0, 0), as no illusion is perceived when the mountain slope is 0.

Experiments 1 and 2: visual cues

In Experiment 1, the average perceived tilt of the buildings relative to gravitational vertical was 19.4° at night, increasing linearly as a function of the mountain slope.
Fig. 2. Illustration of and results for Experiments 1, 2, 6, 7, and 8. The scatter plots (with best-fitting regression lines) show the perceived tilt of nearby skyscrapers as a function of the mountain slope, with data from each observer shown separately for each condition. In Experiment 1 (a), observers viewed buildings while sitting in a reclining position at both day and night (only results for nighttime are shown here to provide a baseline comparison for all other experiments, which were conducted only at night). In Experiment 2 (b), observers viewed buildings through a box with a round, 10-cm opening in front. In Experiment 6 (c), observers sat with a compensating wedge behind their back, under their buttocks, and underneath their feet to position them at gravitational vertical to compensate for the mountain slope. In Experiment 7 (d), observers stood freely on the floor of the tram. In Experiment 8 (e), observers stood on a foot wedge while wearing a box with a round, 10-cm viewing hole over their head. The data points and dashed regression lines from (a) are overlaid on the results shown in (b) through (e), whereas the solid regression lines and open circles in (b) through (e) refer to the experimental condition in each panel.
The Hong Kong Peak Tram Tilt Illusion

(Figs. 1c and 2a). The best-fitting regression line showed that the perceived tilt \( \theta \) matched the slope of travel \( \alpha \), which suggests complete adaptation of the subjective vertical away from gravity (slope\(_{\text{night}} \) = 1.030, \( r = .74 \)), \( t(240) = 17.09, p < .0001 \). In comparison, the slope for measurements obtained in daylight was significantly reduced (slope\(_{\text{day}} \) = 0.721, \( r = .64 \); slope\(_{\text{night}} \) – slope\(_{\text{day}} \) = 0.309), \( t(466) = 9.231, p < .001 \), which suggests an undercompensation of backward pitch position (the average illusory size was 12.85°).

Given that all other factors remained identical, the greater perceived tilt during nighttime was likely to have originated from the visual system. It may have resulted from the relative absence of visual-orientation cues outside the tram window, which are normally available at daytime but are progressively lost as it gets dark. Alternatively, it may be attributable to a heightened sense of enclosure during nighttime tram rides, which strengthens the impact of features within the compartment, such as oblique window frames, beams, the floor, and lighting fixtures. These features all provide strong contextual references of a tilted frame in the immediate environment, and these references compete with the visual orientation cues in the outside world. As a consequence, the true vertical given by the high rises would be expected to appear tilted in the opposite direction, similar to the rod-and-frame effect, in which an upright line appears tilted when presented in a tilted frame (Asch & Witkin, 1948).

When all interior references were masked by the box over observers' heads in Experiment 2, the apparent tilt of the buildings was significantly reduced by about 7° (slope = 0.66, \( r = .44 \); change in slope relative to baseline = 0.373), \( t(382) = 9.611, p < .001 \) (Fig. 2b); this finding was consistent with our prediction that the greater prominence of misleading references inside the compartment caused a greater perceived tilt. However, it did not correct the erroneous perceived tilt completely, which implies that other factors are involved in this misperception.

Experiments 3, 4, 5, 6, and 7: nonvisual cues

Using each wedge by itself (Experiments 3–5) produced only a small effect, if any, on perceived tilt. However, when the three compensating wedges were used together in Experiment 6, the perceived tilt (14.6°) was significantly reduced by about three quarters relative to the baseline condition (slope = 0.79, \( r = .56 \); change in slope relative to baseline = 0.24), \( t(430) = 6.79, p < .001 \) (Fig. 2c). Because we were unable to build a wedge that dynamically changed its compensating angle as the tram moved, the choice of the fixed 18.7° wedge was the most parsimonious solution we could devise; however, it inevitably overcompensated (at slopes less than 18.7°) and undercompensated (at slopes more than 18.7°) during some parts of the tram ride. We separated the overcompensated and undercompensated zones for analysis and did not find that these had any systematic effects on illusory tilt. Our findings demonstrate that signals from all the senses must be consonant with each other to abolish the tilt illusion. Vertically orienting the head without vertically orienting the body is not sufficient.

When observers stood up on the tilted floor of the tram and maintained a gravitationally vertical posture in Experiment 7, comparable with an inverted pendulum, the average perceived tilt (11.16°) was reduced by about 40% relative to the baseline condition (slope = 0.61, \( r = .42 \); change in slope relative to baseline = 0.42), \( t(445) = 12.72, p < .001 \) (Fig. 2d). The reduction was greater than that in Experiment 6, in which all three wedges were used. In both Experiments 6 and 7, the observer's upper body and head were restored to the gravitational vertical; the additional reduction of the illusion while standing suggests that the muscular force exerted by the legs during active balancing is a major factor in defining the true vertical.

Up to this point, we found that (a) the Hong Kong Peak Tram illusion was still present during the day, although much weaker than during the night; (b) masking contextual references inside the tram compartment (i.e., by putting a box over the head) reduced the illusion; (c) three compensating wedges when applied together reduced the illusion and had a greater effect than when each wedge was used alone; and (d) an active standing position caused the greatest reduction of all.

Experiment 8: combined visual and nonvisual cues

None of the manipulations in the previous experiments fully cancelled the skyscrapers’ apparent tilt. In the final experiment, observers were therefore asked to stand on a wedge to render their feet horizontal while viewing the environment through the box used in Experiment 2. As a result, in Experiment 8, the illusory tilt disappeared almost entirely for 3 of our 4 observers (slope = 0.195, \( r = .35 \); change in slope relative to baseline = 0.70), \( t(291) = 22.02, p < .0001 \) (Fig. 2e). We conclude that the combination of having observers stand upright on a wedge and occluding the oblique tram window frame and other references from the interior restores the percept of the vertical by causing observers to draw on the combined information from the vestibular, tactile, proprioceptive, and visual senses. This finding convincingly demonstrates the crucial role of congruent sensory signals for the perception of verticality. This is consistent with findings of neurophysiological studies, which have shown that during sensory integration, animal and human subjects dynamically weigh each
sensory cue based on its spatial and temporal congruency with other sensory inputs (Avillac, Hamed, & Duhamel, 2007; Festsch, Turner, DeAngelis, & Angelaki, 2009; Gu, Angelaki, & De Angelis, 2008; Meredith, Nemitz, & Stein, 1987; Meredith & Stein, 1996) in a manner similar to that of an ideal observer (Alais & Burr, 2004; Ernst & Banks, 2002; Helbig & Ernst, 2007).

The Hong Kong Peak Tram illusion is multisensory in nature

When results from all the experiments were combined, they were in line with a weighted linear model, which suggests a comparator that receives, weights, and possibly modulates the contributing sensory signals during sensory integration (Ma, Beck, Latham, & Pouget, 2006; Ohshiro, Angelaki, & DeAngelis, 2011). Our conclusions are reinforced by observations on a Japanese mountain tram going up to the Hakone resort, in which no tilt illusion is ever experienced. This is presumably because in the Hakone mountain tram, the window frames are carefully set to discount the mountain slope. As a result, there are no misleading visual references from the tram compartment. In addition, observers sit on chairs that adjust to align the observers’ posture to the gravitational horizontal and vertical at all times (similar to the hydraulically adjusted compartments of the Nordketten mountain tram in Innsbruck, Austria).

Discussion

The Hong Kong Peak Tram illusion is novel in its unique combination of an unusual environment with tall nearby buildings and a transportation vehicle that generates large dynamic body pitches. This illusion is robust, as neither experience from frequent exposure nor better knowledge about the illusion eliminates the perceived tilt—observers in this study still experienced the illusion vividly after some 100 tram trips. At the same time, it is puzzling that the tilt is perceived only for the high rises next to the tramline but not with distant buildings farther away. In fact, one skyscraper is seen several times from near and far as the tram goes around a bend, and this same building appears affected or unaffected by the illusion depending on whether it is closer or farther from the tram. The interplay between perceived tilt, distance, object size, and illusory effect should be explored in future investigations.

Our results suggest that a cross-modal verticality system is responsible for the misperception of verticality experienced on the Hong Kong Peak Tram, which differs from well-known past demonstrations (e.g., the tilted room and haunted swing) that are primarily visually induced. Studies on spatial orientation have indicated that people rely on both allocentric (i.e., gravitational direction) and egocentric (i.e., observers’ body orientation) reference frames (Howard, 1982; Rock, 1990; Wade, 1992). These frames are both multisensory in nature (Luyat & Gentaz, 2002), and how each sensory stage interacts and is integrated in verticality perception is an important question for future research.

These experimentally controlled observations from an enclosed tram operating in natural surroundings open a new avenue for the investigation of the phenomenal, neurobiological, and computational aspects of cross-modal sensory integration. They also point toward the influence of contextual cues in enhancing the environmental relevance of daily perceptions. In contrast to the artificially reduced stimuli commonly used in the laboratory, we showed under field conditions that gravitational signals from the senses become integrated in a complex manner with the real environment.

Author Contributions

All authors participated in designing the study, collecting and analyzing the data, and writing the manuscript. All authors approved the final version of the manuscript for submission.

Acknowledgments

We are indebted to the Hong Kong Peak Tram Company for permitting us to conduct our experiments aboard their trams during regular hours of operation. We thank K. F. Tam and Jason Carlow for their assistance with our measurement devices. We are grateful to Pinghui Chiu, Chufu Zhong, and Anna Ho for assistance with data collection and to Matt Oxner for data analysis and technical help.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

The study was supported by grants from the Hong Kong Grant Research Council and the University of Hong Kong Seed Funding Programme for Basic Research to Chia-huei Tseng, and by awards from the Serena Yang Educational Fund and the Deutscher Akademischer Austauschdienst (German Academic Exchange Council) and National Science Council of Taiwan to Lothar Spillmann.

References


