Effect of Scenario on Perceptual Sensitivity to Errors in Animation

Paul S. A. Reitsma*
Carol O'Sullivan†
Trinity College Dublin

Abstract

A deeper understanding of what makes animation perceptually plausible would benefit a number of applications, such as approximate collision detection and goal-directed animation. In a series of psychophysical experiments, we examine how measurements of perceptual sensitivity in realistic physical simulations compare to similar measurements done in more abstract settings. We find that participant tolerance for certain types of errors is significantly higher in a realistic snooker scenario than in the abstract test settings previously used to examine those errors. By contrast, we find there is no difference in tolerance vs. the abstract setting when those errors are displayed in a realistic but more neutral environment. Additionally, we examine the interaction of auditory and visual cues in determining participant sensitivity to spatiotemporal errors in rigid body collisions. We find that participants are predominantly affected by visual cues. Finally, we find that tolerance for spatial gaps during collision events is constant for a wide range of viewing angles if the effect of foreshortening and occlusion caused by the viewing angle is taken into account.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—;

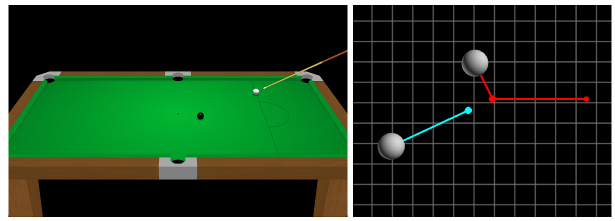
Keywords: psychophysics, animation, perception

1 Introduction

There are several reasons why an animation application might deviate from physically-correct rigid body motion, such as computational savings from approximate collision detection [O'Sullivan and Dingliana 2001] or achieving a particular animation result [Barzel et al. 1996]. Deviating too far from physical correctness, however, can lower the perceived naturalness of the animation [O'Sullivan et al. 2003; Reitsma and Pollard 2003]. Furthermore, it has been suggested that the increased realism of modern rendering techniques could additionally constrain the acceptable range of deviations, due to the increased sensitivity of users to more detailed displays [Stappers and Waller 1993] and mismatched quality levels between animation and rendering.

In is unknown, however, whether people tend to be more or less sensitive to errors in more realistic environments. Can particular aspects of a scenario, such as audio or textures, be manipulated in order to raise or lower user tolerance to errors in the animation? Or, in contrast, do users typically notice errors in animation regardless of the context offered by a scene?

*e-mail: psar@cs.cmu.edu
†e-mail: carol.osullivan@cs.tcd.ie



(a) (b)

Figure 1: *(a) Our snooker scenario. (b) The test scenario of O'Sullivan et al.*

We examine user sensitivity to angular and spatio-temporal errors applied to physically simulated rigid body dynamics in a realistic environment (Figure 1(a)). We compare our results to previous studies of similar errors conducted using abstract stimuli (Figure 1(b)), and examine potential causes for the observed differences. In addition, we examine the relative importance of visual and auditory cues for spatiotemporal errors during rigid body collisions.

We find that the choice of scenario used for testing can significantly bias user sensitivity to angular distortions, but that the addition of a high-contrast texture to provide rotational information does not affect user sensitivity to such distortions. Sensitivity to brief delays in the animation at the time of collision appears to be invariant to scenario and to the timing of audio from that collision; indeed, we find no evidence that audio cues affect user sensitivity to any types of errors. Finally, we find that the overhead view used in many experiments results in equivalent or more conservative error tolerance thresholds than alternative viewing angles, and user sensitivity appears to change slowly with modest deviations from the overhead angle.

2 Background

A number of researchers have suggested techniques for exploiting approximate physics in animations, especially approximate collisions. O'Sullivan and Dingliana [2001] examined perceptual thresholds for approximating collisions to reduce computational complexity, while Barzel et al. [1996], Chenney and Forsyth [2000], Popović et al. [2000], and Twigg and James [2007] used approximate collisions to achieve plausible goal-directed animations. One of our aims is to provide guidance on perceptual sensitivity to various types of errors in order to allow tools such as these to be used more effectively.

The interactions of a small number of simple objects have been studied for decades [Michotte 1963; Cohen 1964; Stappers and Waller 1993; Kaiser and Proffitt 1987; O'Sullivan and Dingliana 2001; O'Sullivan et al. 2003]. We draw on the experiments of Kaiser and Proffitt [1987] and O'Sullivan et al. [2003] as the starting point for our experiments. Kaiser and Proffitt examined user sensitivity to a variety of errors applied to the collision of two circular bodies in an abstract 2D environment, and included a simple model of friction (constant deceleration) in their experiments. O'Sullivan et al. extended the examination of these errors to an abstract 3D environment, but did not consider friction. We examine sensitivity to these errors in a physical simulator, providing a visually realistic environment with physically-correct dynamics.

Research has demonstrated increased user sensitivity to motion displayed on richer and more detailed models, including a fountain with varying numbers of water droplets [Stappers and Waller 1993] and more or less realistic humanoid characters [Hodgins et al. 1998]. Similarly, Oesker et al. [2000] reported more detailed and realistic animation of humanoid football players resulted in more accurate user discrimination of relative skill. We examine whether there is a similar link between realism or detail and user sensitivity in the case of rigid body animation errors.

Recent multimodal perceptual research shows that visual and auditory motion cues can potentially interact. Alais and Burr [2004] report a small increase in sensitivity to bimodal motion (i.e., simultaneous apparent motion of a sound source and visual stimulus), but no directional effect (i.e., visual and auditory motion in the same direction is no more detectable than visual and auditory motion in opposite directions). Their results suggest that visual and auditory cues may be processed independently and then combined at the participant's decision stage. However, auditory and visual cues can also interfere in some contexts. The Metzger illusion [Metzger 1934] involves two dots, one moving left to right and the other moving right to left along the same level. When the two dots meet their interaction is ambiguous, as they could be perceived as either bouncing off or passing through each other; however, an auditory cue played at the moment the balls touch results in a consistent perception that the balls are bouncing off of each other[Sekular et al. 1997]. McGurk and MacDonald [1976] found that given a video of a person saying one phoneme which had been dubbed over with a different phoneme, participants perceived a third phoneme intermediate between those presented by the two stimuli. Accordingly, we investigate whether a similar multimodal interference occurs in the perception of spatiotemporal errors in animated motion.

3 Experimental Setup

As the testbed for our experiments we use a physically-simulated snooker environment (Figure 2(a)). A full physical simulation was used so as to correctly model factors such as sliding and rolling friction; additionally, it offers the ability to cleanly and correctly make changes to any aspect of the inertia tensors of the simulated balls.

We selected snooker as the subject of the simulation primarily due to its innate similarity to the abstract environments used in many previous experiments. Additionally, however, snooker has the benefit of being a familiar and easily-understood scenario, which we surmised would help strengthen the sense of realism we wished to examine.

For all experiments, ball placement was handled identically, with most aspects randomized in order to prevent systematic bias. A target pocket on the table was selected randomly, and the target ball was placed 65cm from the pocket. Its precise position at that distance was randomly chosen such that its eventual trajectory would take it towards that pocket without being at too shallow an angle (20^o) with respect to the sides of the table. Furthermore, the placement was determined such that there was a 50% chance each of the ball sinking into the pocket or missing by a random angle up to 10^o. The cueball was then placed 65cm from the target ball with its angle randomly chosen such that it would strike the target ball from either left or right (as chosen) at a 20—40^o angle, and such that the target ball would follow its prescribed trajectory. (No effects of target pocket, left/right side of table, or top/bottom of table were found for our experiments.)

The simulation started with both balls at rest. A short animation of a cue (or other striking object in the non-snooker scenario) withdrawing and then contacting the cueball was followed by the cueball accelerating to a velocity of $1.8m/s$, after which all further motion was physically simulated. Simulation was terminated 2s after the first collision between the balls, at which point a response screen was overlaid.

The stimuli were shown on a 51cm by 32cm display. participants sat approximately 90cm from the screen, so the display occupied approximately 32^o of their visual field. A regulation snooker table (3.6m plus sides) extended across the width of the screen, meaning 1cm of screen distance corresponded to approximately 7.6 cm of simulation distance, and 12cm of simulation distance corresponded to approximately 1^o of visual field. Each snooker ball had a diameter of 1.0 cm on the screen. Participants wore headphones, and were instructed that in some instances the simulation included audio. Participants were instructed to take into account all information from the simulation to determine whether the event was realistic or whether an error was present. Responses were registered by using the left and right index finger triggers of a gamepad to select "yes" or "no" when prompted by on-screen cues.

4 Study 1: Effect of Scenario Realism on User Sensitivity

Our first study examined user sensitivity to post-collision angular distortions in our snooker simulator (Figure 2(a)). We examined four cases:

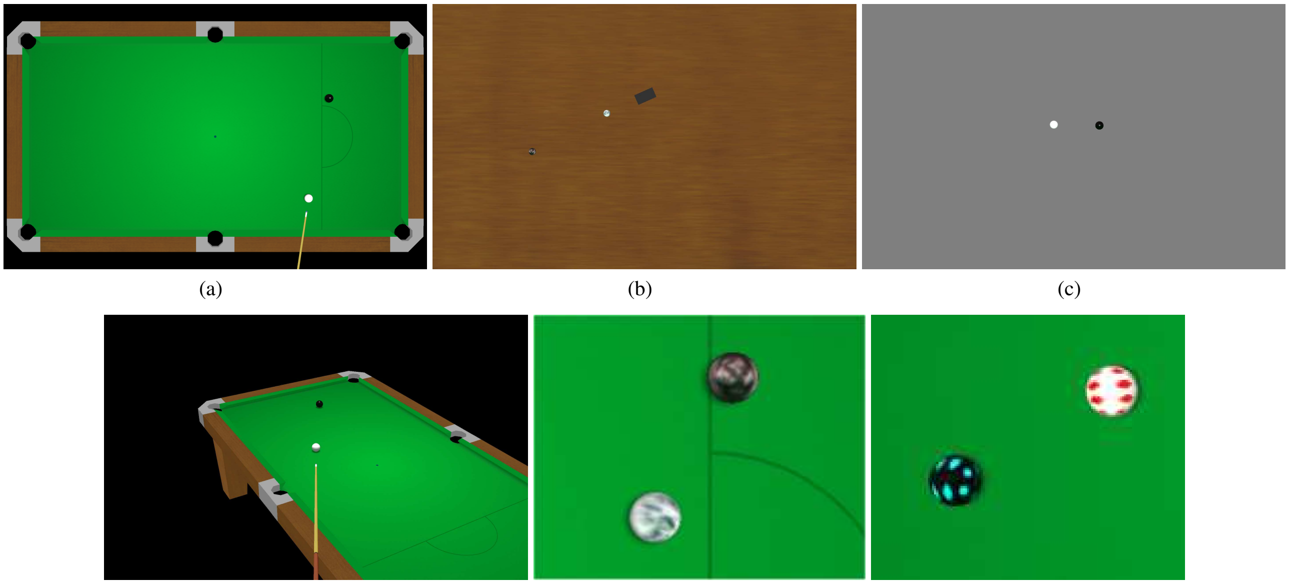
1. Expansion of target ball post-collision angle, clockwise.
2. Expansion of target ball post-collision angle, counterclockwise.
3. Expansion of cueball post-collision angle, clockwise.
4. Expansion of cueball post-collision angle, counterclockwise.

Our hypothesis was that the results would not differ from those reported by O'Sullivan et al. [2003].

Each of these cases was evaluated using randomly-interleaved ascending and descending staircases [Cornsweet 1962; Levitt 1971] with eight reversals. Staircase methods are adaptive tests which rapidly home in on a participant's perceptual threshold, which can improve the efficiency of studies. In addition, combining ascending and descending staircases helps to avoid misinterpreting results due to guessing, as near-random responses will tend to result in the ascending staircase converging to a substantially lower value than the descending staircase, indicating low reliability.

Alterations to the post-collision angle of the cueball and the target ball were presented in separate blocks, for a total of two blocks of four staircases each. A cumulative normal distribution function (ogive) was fitted to the responses of each participant for each of the four experimental conditions, allowing the participant's *Point of Subjective Equality* (PSE; the error magnitude at which they would be 50% likely to consider a motion as having an error) and *Just Noticeable Difference* (JND; the difference between the error magnitudes required to elicit 50% and 75% rejection rates) to be computed. Added errors were capped at 120^o in order to prevent wrapping around 360^o or settling into local minima, and participants who persistently responded that motions with maximum error were realistic were assigned a value of 120^o for that error condition.

There were 19 volunteers in this study: 14 male and 5 female staff and students, aged 13 to 46 (mean 27). All participants had normal



(a) (b) (c) (d) (e) (f)

Figure 2: *Test environments used in our experiments. (a) Base snooker scenario. Cueball is white, target ball is black. (b) The non-snooker environment used in experiment 2. (c) The neutral environment used in experiment 3. (d) The snooker environment seen from a 60^o angle. (e) A closeup of the texture applied to the balls in experiment 2. (f) A closeup of the texture applied to the balls in experiment 3.*

Ball	Direction	Mean PSE	Mean JND
Target	Clockwise	$46.7^o \pm 3.4^o$	$16.9^o \pm 4.0^o$
Target	CCW	$46.4^o \pm 3.8^o$	$17.4^o \pm 3.8^o$
Cueball	Clockwise	$82.8^o \pm 6.8^o$	$57.7^o \pm 11.7^o$
Cueball	CCW	$88.1^o \pm 8.2^o$	$58.0^o \pm 11.2^o$

Table 1: *Results for our first study. Point of subjective equality (PSE) is the magnitude of error where a participant was 50% likely to notice and remark on it. Just noticeable difference (JND) is the additional magnitude of error required to change from 50% to 75% rejection threshold. Values are given as mean $\pm$ standard error of the mean (SEM).*

or corrected-to-normal vision, were naive as to the types of errors being examined, and were given the same set of instructions.

4.1 Results

The results of these experiments are shown in Table 1. Throughout, we applied t-tests for between-study comparisons and paired t-tests for within-study comparisons. As in the simpler and less realistic scenario of O'Sullivan et al.[2003], no difference was found between clockwise and counterclockwise expansion; however, participants were in general more tolerant of errors in our experiment. Our mean point of subjective equality (PSE) for the target ball was $46.5^o \pm 2.5^o$ as compared to 30^o found by O'Sullivan et al.; similarly, our mean PSE for the cueball was $85.4^o \pm 5.3^o$, which again differs substantially from O'Sullivan et al.'s mean PSE for the striking ball of 60^o.

Insight into this difference can be gleaned from post-study responses and from the pattern of results for which staircase convergence was poor. Some participants, typically those who rated their familiarity with snooker as high on the post-study questionnaire, reported that some of the collisions they saw were possible, but only with high levels of spin on the cueball; some even reported that they had been strongly influenced by their belief that they could personally have made many of those shots.

Similarly, while most data was well-approximated by an ogive, there were two characteristic types of non-convergence. The first is when the ascending and descending staircases for a particular error condition converged to different values; the second is when a participant consistently answered "realistic" to motions with maximum error, and hence did not converge to a fixed value. While 6 of 11 participants who reported low experience with snooker had the first type of non-convergence on one or both cueball error treatments, none of the 8 participants who reported high experience with snooker converged to two different values in that manner. By contrast, none of the 11 low-experience participants failed to converge in the second manner, whereas 3 of 8 high-experience participants had type 2 non-convergence. This distribution is statistically significant ($\chi^2 = 12.0, P = 0.0025$), suggesting that the context evoked by a realistic scenario can strongly affect the responses of a participant, and that this effect might explain the lower sensitivity to errors seen in our experiment as compared to O'Sullivan et al. However, other differences between the scenarios (ball size and velocity, presence of friction, etc.) might also account for the difference.

5 Study 2: Effect of Scenario Context and Audio Cues

The goal of our second study was to explore the potential for a scenario to provide context that would bias a participant's expectations and hence their tolerance for errors. When examining the effect of scenario and realism, a natural question was what effect the additional information and verisimilitude provided by realistic audio cues would have. Our hypotheses were:

- Participants would be more tolerant of errors in a realistic snooker scenario.
- Audio cues would lower participant tolerance of errors in general, but would raise tolerance if the audio and visual cues conflicted.

Error	Scenario	Mean PSE	Mean JND
Gap	Snooker	$5.18mm \pm 0.73$	$2.29mm \pm 0.78$
	Non-snooker	$5.79mm \pm 0.88$	$3.94mm \pm 0.92$
Cueball	Snooker	$87.5^o \pm 10.4^o$	$40.4^o \pm 8.3^o$
	Non-snooker	$53.1^o \pm 12.2^o$	$47.5^o \pm 12.8^o$
Target	Snooker	$39.4^o \pm 4.9^o$	$14.5^o \pm 4.0^o$
	Non-snooker	$33.6^o \pm 4.6^o$	$13.7^o \pm 2.9^o$

Table 2: *Results for our scenario effect study. Point of subjective equality (PSE) is the magnitude of error where a participant was 50% likely to notice and remark on it. Just noticeable difference (JND) is the additional magnitude of error required to change from 50% to 75% rejection threshold. Values are given as mean ± standard error of the mean (SEM). Gap distances are given in terms of on-screen distance; 16mm is approximately one degree of a participant's field of view.*

There were 16 volunteers in this study: 13 male and 3 female staff and students, aged 20 to 37 (mean 26). All participants had normal or corrected-to-normal vision, were naive as to the types of errors being examined, and were given the same set of instructions with the added instruction that they were to judge the simulation on its own merits, rather than trying to second-guess whether something they noticed was intentional or not. This additional instruction was added after some participants in pretests reported that they consistently ignored delay errors that they knew were present, under the assumption that such errors were not intended to be part of the test. Trials were presented in random order.

5.1 Scenario Effect on Tolerance

In order to test our hypothesis that participants' sensitivity to errors was being affected by preconceived ideas regarding snooker, we prepared a second scenario (Figure 2(b)) which used identical motion and viewpoint as the snooker scenario, but was rendered to look like marble spheres rolling on a wooden plane rather than being a snooker game.

We tested *gap errors* as well as angular distortion errors in these two environments in order to see if the effects applied to different types of errors. A gap error resulted in the first collision between the two balls occurring as if the cueball had its radius increased by the magnitude of the error.

Angular distortion errors were presented as in the first experiment, and both types were evaluated using matched ascending and descending staircases. Maximum error value for gap errors was 250mm of physical distance in the simulation, which corresponded to approximately 33mm of distance on the participant's screen. Balls were given identical textures in all error conditions.

5.1.1 Results

Table 2 shows the result of scenario on tolerance for error. The main result is that scenario had a strong effect on participant tolerance for some types of errors.

Participants' tolerance for angular errors in the non-snooker scenario was substantially lower than in the snooker scenario (cueball: paired-$t = 2.46, P = 0.02$; target ball: paired-$t = 1.73, P = 0.05$), and was not substantially different from the values reported in the work of O'Sullivan et al. [2003] (cueball: $53^o \pm 12^o$ vs. 60^o; target ball: $34^o \pm 5^o$ vs. 35^o).

By contrast, there was no difference in participant tolerance to gap errors in the snooker vs. non-snooker scenarios ($5.2mm \pm 0.7mm$ vs. $5.8mm \pm 0.9mm$). While participants were more tolerant of gap errors in our experiment than in O'Sullivan et al. (mean PSE 5.5mm vs. 0.7mm), the differing size and speed of the objects makes it impossible to compare the tasks directly.

Error	Sound	Mean PSE	Mean JND
Delay	Silent	$62.0ms \pm 10.5$	$28.2ms \pm 9.6$
	Early	$62.1ms \pm 9.4$	$30.8ms \pm 7.7$
	Late	$62.7ms \pm 10.8$	$39.3ms \pm 13.8$
	Sound	$61.1ms \pm 42.6$	$120.5ms \pm 122.5$
Gap	Silent	$6.17mm \pm 0.73$	$2.15mm \pm 0.54$
	Audio	$6.49mm \pm 0.75$	$2.17mm \pm 0.35$
Angle	Audio	$46.0^o \pm 5.9$	$12.4^o \pm 2.4$
	Silent	$46.5^o \pm 3.6$	$17.1^o \pm 3.9$

Table 3: *Results for our audio cue study. Note that data for angular distortions with no audio is from study 1. Point of subjective equality (PSE) is the magnitude of error where a participant was 50% likely to notice and remark on it. Just noticeable difference (JND) is the additional magnitude of error required to change from 50% to 75% rejection threshold. Values are given as mean ± standard error of the mean (SEM). Gap distances are given in terms of on-screen distance; 16mm is approximately one degree of a participant's field of view.*

5.2 Effect of Audio Cues

Realistic audio cues added to the simulation were for collision events (cue/ball, ball/ball, ball/sidewall, and ball/pocket).

Our hypothesis was that the additional information given by audio cues would increase participant sensitivity to gap and angular distortion errors. In addition, we examined whether user perception of *delay errors* could be altered by playing audio cues before the delay or after the delay. A delay error of Nms caused the simulation to pause for the indicated duration the moment the two balls touched for the first time. For trials where audio output was enabled, the sound of the collision was played either before pausing the simulation, referred to as *early delay*, or after the end of the pause (*late delay*). Finally, we examined the case where the animation was not paused at all, but the sound of collision was delayed (*sound delay*). We hypothesized that this would have a similar effect on perceived realism as a delay in the animation. Maximum error value was 250ms for the early and late delay cases, and 500ms for sound delay.

5.2.1 Results

Table 3 shows the results of this experiment. We found that audio cues had no effect on participant sensitivity.

Participants attended almost exclusively to visual cues, and audio cues had no significant effect on any error treatment. In particular, there was no effect on participant sensitivity to delay errors regardless of whether the collision sound was played early, late, or not at all. Moreover, data for the majority of participants did not converge well for the "sound delay" error condition, suggesting they were not significantly attending to the audio cue. Table 3 reports data on the "sound delay" error condition only for those participants whose data converged well, the effects of which are considered in Section 7.

We note also that participant tolerance of delay errors did not differ between our study and prior work (mean PSE 62ms vs. 60ms).

6 Study 3: Characterizing User Sensitivity

Our goal for the third experiment was to examine the scenario effect in greater detail by using a different study design. In particular, we wanted to examine whether response bias played a role in the observed differences in participant tolerance for errors between scenarios.

Additionally, prior work and pretests of our own had shown that participant tolerance for gap errors increases as viewing angle from the vertical increases. For a fixed size of gap, however, the *visible gap*, or apparent size of the physical gap, decreases with increased viewing angle, both due to foreshortening and due to occlusion by the nearer ball, and it would be useful to know whether this quantity accurately predicts participant tolerance for gap errors.

Our hypotheses were:

- Participants would have lower sensitivity to errors in the snooker environment than in a more abstract environment.
- Balls with high-contrast textures would lower participant tolerance to angular distortions.
- Participant tolerance to gap errors at different viewing angles is determined by the visible gap at that angle.

There were 15 volunteers in this study: 10 male and 5 female staff and students, aged 19 to 43 (mean 25). All participants had normal or corrected-to-normal vision, were naive as to the types of errors being examined, and were given the same set of instructions as for the previous experiments, including the instruction to not second-guess which observations were intended to be part of the experiment.

6.1 Scenario and Texture

Angular distortions to the target ball were tested in the snooker scenario (Figure 2(a)) and in a frictionless neutral scenario (Figure 2(c)), with the neutral scenario appearing first in order to prevent participants from associating it with snooker. Initial velocity in the frictionless scenario was lowered in order to make post-collision velocities similar between the two scenarios.

In order to evaluate the scenarios for participant bias, we adopted a repeated measures design, which allowed a detection-theoretic analysis to be performed. As noted by Reitsma and Pollard [2003], detection theory [Macmillan and Creelman 1991] can be used to derive a bias-independent measure of a user's ability to detect errors in an animated motion. The method takes into account the difference between how frequently the subject correctly labelled a motion as containing an error (hit rate H) and how frequently the subject incorrectly labelled an unchanged motion as containing an error (false alarm rate F). A subject's *sensitivity* (d) to errors iscomputed as:

$$d = z(H) - z(F) \tag{1}$$

where z is the inverse of the normal distribution function. For example, a hit rate of 50% and a false alarm rate of 16% corresponds to a sensitivity of 1.0, as does a hit rate of 30% coupled with a false alarm rate of 6%. These two examples of how to obtain a sensitivity of 1.0 demonstrate the bias-independent nature of detection theory: as sensitivity is computed based on the relative distribution of the participant's responses rather than on the raw distribution, factors which will systematically bias the responses, such as participant reaction to the level of realism with which the scenario is rendered, are automatically factored out.

Five levels of angular distortion (15^o—75^o) were added to collisions, with each error treatment being repeated three times. Identical collisions without distortion were added to balance, resulting in a total of 30 motions. Each of these motions was displayed with the balls textured to appear as a snooker ball or as in Figure 2(f), for a total of 60 trials per scenario. As well as computing sensitivity measures per participant and in aggregate, PSEs and JNDs were computed in the same manner as for previous experiments.

Angular Distortion	Aggregate Sensitivity: Snooker Environ Base	Snooker Environ Dots	Neutral Environ Base	Neutral Environ Dots
0^o	0	0	0	0
15^o	0.53	0.04	-0.32	-0.25
30^o	0.98	0.64	0.92	0.61
45^o	1.64	1.56	1.65	1.25
60^o	2.45	2.29	2.15	1.73
75^o	3.43	2.73	2.44	2.19

Table 4: *Aggregate participant sensitivities to angular distortion errors in experiment 3. "Base" denotes trials where the snooker balls had their normal appearance, while "Dots" denotes trials where the snooker balls had a dot texture applied, as in Figure 2(f). Sensitivity was highest in the most realistic scenarios, and lowest in the least realistic.*

6.1.1 Results

Table 4 contains the sensitivity results for this study. We note the surprising result that sensitivity was negative for small angular distortions in the neutral environment, suggesting that participants actually preferred slightly-expanded collision angles to fully-realistic ones.

We found that participant PSEs for angular distortion were again lower in the neutral environment than in the snooker environment, and by approximately the same amount as with the previous experiment ($35.6^o \pm 3.0^o$ vs. $43.8^o \pm 4.4^o$ for the non-textured balls), although the difference was only weakly significant (paired-t = 1.64, P = 0.07). Participants were much more likely, however, to flag a motion as containing an error in the neutral environment (unbalanced ANOVA F(1,1498) = 14.8, P $<$ 0.001). This response bias resulted in nearly triple the false alarm rate in the neutral environment as in the snooker environment (15.6% vs. 5.4%), leading to the unusual situation in which participants were both more tolerant and more sensitive in the snooker environment (mean sensitivity 1.23 ± 0.13 vs. 1.46 ± 0.11).

Contrary to our expectations, we found that texture had no significant effect on participant PSEs or on tendency to report that a motion contained an error. We did find that the textures resulted in slightly lower participant sensitivity to errors (1.25 ± 0.12 vs. 1.45 ± 0.12, paired-t = 2.80, P = 0.01) due to a higher false alarm rate (11.7% vs. 8.8% with no textures; hit rate was 47.5% and 45.6%, resp.).

6.2 Visible Gap

Figure 2(a) shows an overhead view of two snooker balls 650mm apart, and Figure 2(d) shows an identical physical distance between two snooker balls as seen from a viewpoint along the cue at a 60^o angle to the vertical. Due to the differing viewpoints, the same size of physical gap results in a difference visible gap between the two balls.

For an observer looking along the line between the centres of the two balls a viewing angle which is θ degrees from the vertical (overhead position), a physical distance d between the two balls will re-

Viewing Angle	PSE for Physical Gap (mm)	PSE for Visible Gap (mm)
0^o	6.39 ± 0.98	6.39 ± 0.98
20^o	7.44 ± 1.08	6.76 ± 1.01
40^o	9.38 ± 1.33	6.12 ± 1.02
60^o	12.17 ± 1.38	2.59 ± 0.69
80^o	21.58 ± 1.76	0

Table 5: *Mean PSE values for detection of gap errors at different viewing angles from the vertical, along with the amount of gap visible at that angle. Values are given as mean ± standard error of the mean (SEM). Gap distances are given in terms of on-screen distance; 16mm is approximately one degree of a participant's field of view.*

sult in a visible gap between them of:

$$V = cos(\theta) * d - r * (sin(\theta)tan(\theta) - 1 + cos(\theta)) \qquad (2)$$

where r, the radius of the balls, is 26.25mm. At a 60^o viewing angle, a 650mm physical gap corresponds to a visible gap of 298.75mm, a reduction of 54%. These distances in the simulation correspond to approximately 87mm and 40mm, resp., on the screen used to conduct the experiment. Note that this gives a slightly conservative estimate for visible gap, as our viewpoints were aligned with the cue, which was offset by approximately 1^o from the line between the centres of the balls. For our experiment, the difference was around 1%.

Five viewing angles were examined, from 0^o (corresponding to the default overhead view; Figure 2(a)) to 80^o (corresponding to a view along and just above the cue; Figure 2(d)). Distance from the point of collision was the same for all viewing angles, and the PSE for each viewing angle was estimated as the average of the last four reversals of a descending staircase (out of eight).

6.2.1 Results

PSE for visible gap stayed nearly constant for moderate viewing angles (see Table 5), as the physical distance between balls corresponding to the mean PSE increased just fast enough to offset the manner in which increased viewing angle appears to reduce that distance. At steep viewing angles, however, participants detected gap errors with much lower visible gaps.

7 Discussion

Our most directly-applicable finding is that the visible gap between colliding object appears to be an accurate predictor of user tolerance to gap errors for viewing angles up to about 40^o. We note, however, that there appear to be multiple mechanisms by which participants are able to detect the presence of gap errors. Direct observation of gaps appears to dominate for low viewing angles, as participants converged to a mean PSE of about 6mm of visible gap regardless of viewing angle. At steeper viewing angles, and in particular at the 80^o viewing angle where no gap was ever visible between the snooker balls, participants appear to have used a different technique to detect collision gap errors.

We were surprised to find that audio cues had no significant effect for any of the types of error, in contrast to findings that audio and visual cues can reinforce each other in tasks such as motion detection [Alais and Burr 2004]. Indeed, we note that for the purely audio error condition ("sound delay"), only 8 of 16 participants mentioned sound in their post-study questionnaire, and responses were strongly bimodal, with 6 of 16 participants displaying strong convergence (all PSEs under 150ms, mean 59ms) and the other 10 participants displaying weak or no convergence (all PSEs over 400ms or not computable, as compared to a maximum error value of 500ms). One possibility is that most participants considered visual information to be of overriding importance, and largely dismissed audio cues, despite our explicit instruction to consider all information from the simulation. Indeed, in post-study questioning, one participant expressed surprise that the "sound delay" error condition – i.e., where the animation is correct but the sound is delayed – even existed.

Similarly, neither texture we tried had a significant effect on participant tolerance for errors, and there was no significant difference in participant tolerance for gap errors between any of the three scenarios we tested.

By contrast, we found a significant effect of scenario realism on participant response to angular distortions. Participants' tolerance for angular distortions did not differ significantly between the two non-snooker scenarios we examined, and tolerance in those environments was very similar to that reported by O'Sullivan et al. [2003]; tolerance in the realistic snooker environment, however, was substantially higher. At least for the small set of scenarios we examined, abstract or neutral scenarios appeared to offer a conservative estimate of tolerance; however, it is possible that some scenarios will bias participants to be less tolerant of certain errors.

While participant *tolerance* for errors was higher in the realistic snooker environment, participant *sensitivity* to angular distortions was *also* higher, a seemingly-contradictory result. The root cause for both of these differences appears to be a response bias when reasoning about angular distortion errors; participants were systematically more likely to report a trial in a less realistic scenario as containing an error, regardless of whether or not any error was actually present. Indeed, the most realistic-looking scenario (the basic snooker setup) received the lowest false alarm rate (3.6%) and the least realistic scenario (textured balls in the neutral scenario) received the highest (16.7%). One possible explanation is that unrealistic scenarios increase the amount of randomness and noise in user responses; this would account for the lower sensitivity for the unrealistic environment seen in Table 4. A second explanation is simply that participants were biased by the scenario, and had a greater tendency to view motions in realistic scenarios as being realistic, and conversely motions in unrealistic scenarios as being unrealistic. We note that this response bias appeared to cut across participant level of experience with physics or snooker. The exception is that participants reporting snooker experience displayed equal sensitivity to angular distortions in both the snooker and non-snooker environments, whereas participants with no such experience displayed substantially higher sensitivity in the realistic snooker environment; however, both groups displayed higher tolerance for errors in the realistic environment.

We note that this tendency of more realistic portrayals to result in higher participant sensitivity to some types of of error and not others has been seen for other types of motion (e.g., [Reitsma et al. 2008], and wonder if there might be a consistent pattern to which types of errors are affected. One intuitive, although speculative, possibility is that errors which are local in nature and detected by direct observation (e.g., collision gaps or velocity spikes) would tend to be less affected than errors which are more global in nature, and are more commonly detected by inference from their effect on the overall character of the motion than by direct observation (e.g., angular distortions or changes to gravity). It is possible, however, that particular scenarios could bias other types of errors, even those such as delay which have so far shown no change across scenario types. For example, deformable objects are known to have highly inelastic collisions, and hence a short delay might potentially be confused with object deformation. Finding such a pattern, if one exists, could prove useful in determining to what extent results from certain types of psychophysical experiments can be generalized beyond the scenarios in which the experiments were conducted.

8 Conclusions

In this paper we find that the choice of scenario can strongly influence people's decisions on some types of errors. Tolerance to many commonly-measured errors appears to be consistent between a neutral environment and the type of abstract setup use for some psychophysical studies. We note, however, that choice of scenario can systematically bias the decisions of users regarding certain types of errors, and that different measures of how users are affected by errors may change in different ways due to such biases. Examining both tolerance and sensitivity to errors can help to uncover biases such as these.

We find that neither audio corresponding to the collision event nor the addition of a texture which provided information on object rotation had a significant effect on participant tolerance of errors. We note, however, that in our experiments an unrealistic texture appeared to bias participants to respond more negatively to motions, as did an unrealistic environment. More research is needed to determine how pervasive this type of bias may be.

Finally, we find that visible distance can accurately predict participant sensitivity to gap errors up to angles of 40^o, potentially allowing greater flexibility in tasks such as approximate collision detection.

Acknowledgements

The authors would like to thank Richard Lee for providing the code to his snooker simulator.

References

ALAIS, D., AND BURR, D. 2004. No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research, 19(2)*, 185–194.

BARZEL, R., HUGHES, J. F., AND WOOD, D. N. 1996. Plausible motion simluations for computer graphics animation. In *Proc. Eurographics Workshop on Computer Animation and Simulation*.

CHENNEY, S., AND FORSYTH, D. A. 2000. Sampling plausible solutions to multi-body constraint problems. In *Proceedings of SIGGRAPH 2000*, 219–228.

COHEN, R. L. 1964. Problems in motion perception. *Uppsals, Sweden: Lundequistska*.

CORNSWEET, T. 1962. The staircase method in psychophysics. *American Journal of Psychology*, 485–491.

HICHEUR, H., VIEILLEDENT, S., RICHARDSON, M., FLASH, T., AND BERTHOZ, A. 2005. Velocity and curvature in human locomotion along complex curved paths: A comparison with hand movements. In *Experimental Brain Research*, no. 162, 145–154.

HODGINS, J. K., O'BRIEN, J. F., AND TUMBLIN, J. 1998. Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics 4*, 4 (October), 307–316.

KAISER, M. K., AND PROFFITT, D. R. 1987. Observers' sensitivity to dynamic anomalies in collisions. *Perception & Psychophysics, 42(3)*, 275–280.

LEVITT, H. 1971. Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 467–477.

MACMILLAN, N. A., AND CREELMAN, C. D. 1991. *Detection Theory: A User's Guide*. Cambridge University Press, New York.

MCGURK, H., AND MACDONALD, J. 1976. Hearing lips and seeing voices. *Nature 264* (Dec.), 746–748.

METZGER, W. 1934. Beobachtungen über phänomenale identität. *Psychologische Forschung 19*, 1–60.

MICHOTTE, A. 1963. The perception of causality. *Metheun, London.*

OESKER, M., HECHT, H., AND JUNG, B. 2000. Psychological evidence for unconscious processing of detail in real-time animation of multiple characters. *Journal of Visualization and Computer Animation 11*, 105–112.

O'SULLIVAN, C., AND DINGLIANA, J. 2001. Collisions and perception. *ACM Transactions on Graphics 20*, 3 (July), 151–168.

O'SULLIVAN, C., DINGLIANA, J., GIANG, T., AND KAISER, M. K. 2003. Evaluating the visual fidelity of physically based animations. In *ACM Transactions on Graphics 2, 3*, 527–536.

POPOVIĆ, Z. 2000. Controlling physics in realistic character animation. *Communications of the ACM 43*, 7, 50–58.

REITSMA, P. S. A., AND POLLARD, N. S. 2003. Perceptual metrics for character animation: Sensitivity to errors in ballistic motion. *ACM Transactions on Graphics 22*, 3 (July), 537–542.

REITSMA, P. S. A., ANDREWS, J., AND POLLARD, N. S. 2008. Effect of character animacy and preparatory motion on perceptual magnitude of errors in ballistic motion. In *Computer Graphics Forum (Eurographics)*, vol. 27(2), 201–210.

SEKULAR, R., SEKULAR, A., AND LAU, R. 1997. Sound alters visual motion perception. *Nature, v.385, pp.308*.

STAPPERS, P. J., AND WALLER, P. E. 1993. Using the free fall of objects under gravity for visual depth estimation. *Bulletin of the Psychonomic Society 31*, 2, 125–127.

TWIGG, C. D., AND JAMES, D. L. 2007. Many-worlds browsing for control of multibody dynamics. *ACM Transactions on Graphics 26*, 3.