Humans and other animals are very good at perceiving each other’s movements. In this article, we investigate the role of time-based information in the recognition of individuals from point light biological motion sequences. We report an experiment in which we used an exaggeration technique that changes temporal properties while keeping spatial information constant; differences in the durations of motion segments are exaggerated relative to average values. Participants first learned to recognize six individuals on the basis of a simple, unexaggerated arm movement. Subsequently, they recognized positively exaggerated versions of those movements better than the originals. Absolute duration did not appear to be the critical cue. The results show that time-based cues are used for the recognition of movements and that exaggerating temporal differences improves performance. The results suggest that exaggeration may reflect general principles of how diagnostic information is encoded for recognition in different domains.

Abstract—Humans are very good at perceiving each other’s movements. In this article, we investigate the role of time-based information in the recognition of individuals from point light biological motion sequences. We report an experiment in which we used an exaggeration technique that changes temporal properties while keeping spatial information constant; differences in the durations of motion segments are exaggerated relative to average values. Participants first learned to recognize six individuals on the basis of a simple, unexaggerated arm movement. Subsequently, they recognized positively exaggerated versions of those movements better than the originals. Absolute duration did not appear to be the critical cue. The results show that time-based cues are used for the recognition of movements and that exaggerating temporal differences improves performance. The results suggest that exaggeration may reflect general principles of how diagnostic information is encoded for recognition in different domains.
movement involved in picking up, drinking from, and putting down a glass. These sequences were divided into segments at key frames specifying the start and end of each segment (see Fig. 1). Animators use similar key frames, defining spatial positions in these frames manually while generating intermediate frames automatically (Thom-\(a\)as & Johnson, 1981). As Figure 2a shows, all but one of the key frames that we used fell at local minima on the velocity profile of the wrist, that is, at relatively stationary points between periods of movement. We used the velocity profile of the wrist to determine the key frames, as extremities of this kind are known to be important for the perception of biological motion (Mather et al., 1992). The segments between the key frames correspond to distinct parts of the motion sequence—for example, the first segment corresponds to picking up the glass. The key frame not at a minimum is at a point where the velocity profile changed markedly and corresponds to the onset of actual drinking.

We temporally exaggerated sequences by scaling the durations of the movement segments relative to average values. (See Fig. 2b for the exaggeration of the velocity profile shown in Fig. 2a.) The actor whose movement data the velocity profiles in Figure 2 were derived from picked up, lifted, and drank from the glass in a shorter time than average. In the exaggeration, the durations of these segments became even shorter. Spatial properties like the distance traveled by points between the consecutive key frames shown in Figure 1 remained the same, so spatiotemporal properties, including peak and average velocity of segments, also changed when temporal properties changed. The last two segments of this actor’s movement had durations close to average and were therefore not much affected by exaggeration. (Additional technical details and examples of the animations used are available on the World Wide Web at http://www.psy.gla.ac.uk/harry/ tempexag.html.)

Temporal exaggerations allowed us to test the effect of varying temporal properties of movement while leaving spatial properties constant.

**EXPERIMENT**

In this experiment, participants learned to recognize individuals from their arm movements. The task and stimuli chosen were relatively arbitrary, but had the advantage that the task demands were clearly specified and the information available was controllable. We compared recognition performance for learned and exaggerated sequences, thus determining whether the properties affected by temporal exaggeration are important for this task.

![Fig. 1. Single frames taken from an animation sequence at points corresponding to the key frames shown in Figure 2. The frames represent the maximum spatial extents of the different phases of movement for both unexaggerated and exaggerated sequences. The images are shown in reverse polarity with enlarged dots; they are framed for illustrative purposes. The individual frames represent (a) the start position, (b) picking up the glass, (c) the beginning of drinking, (d) the end of drinking, (e) replacing the glass, and (f) the final position.](image-url)
We tested both time-varying and time-normalized exaggerations. For both types of exaggerations, the relative durations of segments were produced in the same manner—that is, differences from average values were increased by a scaling factor. However, for time-normalized exaggerations, we normalized the overall length of the sequences so that this was always equal to the average value for the actor involved. Comparing these two types of exaggeration provides a test of whether the total duration of movement sequences is itself the critical cue to identifying the actor.

We also varied the level of exaggeration by using different scaling factors. We used two positive levels of exaggeration, with differences increased by factors of +0.5 (+50%) and +1 (+100%); an unexaggerated condition, with differences unchanged (0%); and a negative level of exaggeration, which was scaled by −0.5 (−50%). Positive exaggerations should be better recognized and negative exaggerations worse recognized if exaggeration enhances dimensions of psychological space that are important for recognition by stretching them in a way that facilitates discrimination between individuals.

### Method

#### Participants

Sixteen students from the University of Glasgow took part in the experiment. They were paid for their participation, which took about an hour.

#### Procedure

We presented sequences as QuickTime animations using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). In the initial stage, for each sequence observers saw the first name of the person followed by an original, unexaggerated example of the movement. Three different sequences were shown for each of six individuals, making a total of 18 sequences. Participants did not make any responses during this stage of the experiment.

The second stage of the experiment involved training with feedback. Participants saw 1 of the 18 learned sequences at random and indicated who they thought it was. Responses were made using the keys 1 to 6, each corresponding to a particular actor. A list of which number corresponded to which actor was available throughout the experiment. Automatic feedback indicated whether the response was right or wrong. If the response was wrong, the correct name and number were shown, and that sequence was shown again later in this stage of the experiment. Participants trained until they had correctly identified all 18 sequences twice each. The number of trials to criterion was recorded for each participant.

In the third and final stage of the experiment, participants saw a mixture of the learned sequences and previously unseen, exaggerated sequences in a random order. They responded in the same way as in the second training stage, but there was no feedback and incorrectly identified trials were not repeated. There were 144 trials in this stage (6 actors × 2 types of exaggeration × 4 levels of exaggeration × 3 sequences).

#### Design

The core design was a 2 (exaggeration type: time-varying or time-normalized) × 4 (exaggeration level: −0.5, 0, +0.5, or +0.1) × 6 (actor) within-participants factorial design. Exaggeration levels were the scaling factors, which were applied to the initial differences between the sequence and the grand average.

We had recorded 12 repetitions for each actor when creating the stimuli. In order to make use of all the repetitions and avoid effects due to idiosyncrasies of particular sequences, we produced four versions of the experiment. Each version used a different random triplet of the 12 available sequences for each actor. Version was included as an additional between-participants variable.

Hits and false alarms for each actor in each condition were combined into a single measure of sensitivity, $d_L$, an analogue of $d'$ based on a logistic distribution (Snodgrass & Corwin, 1988). A $d_L$ of zero corresponds to chance performance, and a $d_L$ of 5.4 corresponds to perfect performance (3/3 hits and 0/15 false alarms) in any condition of this experiment.
Temporal Exaggeration

The results confirm that changing temporal properties affects the recognition of identity from biological motion even when the spatial properties remain constant. Exaggerating temporal differences improved performance on a task involving discriminating among people. This suggests that the physical properties that we altered—duration and properties based on duration—are psychologically important for the task of recognizing different individuals from their movements. Exaggerating temporal differences from mean values facilitated recognition although participants had learned unexaggerated versions. Although exaggerations were based on differences in duration, total duration does not appear to be the critical cue, as there was no difference between time-varying and time-normalized exaggerations.

Although theories of object recognition (e.g., Biederman, 1987; Marr & Nishihara, 1978) emphasize the importance of spatial information, primarily shape, it appears that temporally derived information is also useful for recognition. The effect of temporal exaggeration shows that spatial cues, which were available (e.g., the distances between points and the motion path), do not fully determine performance for this task. Participants reported using spatial cues (e.g., the height to which the glass was lifted), but performance was not determined by spatial cues alone, as it was also sensitive to the temporal or spatiotemporal properties of the signal and noise are critical in determining performance on tasks based on the perception of biological motion (Bertenthal & Pinto, 1994; Cutting et al., 1988; Thornton et al., 1998).

There was no effect of whether the absolute duration of the sequence was allowed to vary or was time normalized, ruling out absolute duration as the critical cue. A purely temporal cue like relative duration might still be important, however. For example, the relative durations of the segments of a movement may capture the rhythm of movement and be useful for recognition. Temporal exaggeration might increase individual differences in rhythm and thereby facilitate recognition. However, spatiotemporal cues, such as peak velocity, are also affected by temporal exaggeration, and it may be one or more of these cues that is critical for recognition. Participants mentioned differences in both speed and duration as cues that they had used.

The lack of an effect of time normalizing suggests that whatever whether the results are specific to the perception of biological motion or if they reflect general properties of the perception of moving point stimuli, we conducted a shortened version of the experiment in which all sequences, both learned and tested, were played backward. This manipulation does not affect low-level properties of motion—points still move the same distances in the same time. However, sequences played backward look jerky, perhaps because they disrupt the rhythm of movement and violate implicit knowledge of how people move. In the backward experiment, there was an effect of level of exaggeration, $F(3, 33) = 5.1, p < .05$, but the pattern was very different from that reported for forward sequences. Mean $dL$s (with standard errors in parentheses) were 2.9 (0.3) for the −0.5 exaggeration, 2.9 (0.3) for no exaggeration, 3.2 (0.3) for the +0.5 exaggeration, and 2.6 (0.3) for the +1.0 exaggeration. (These means are higher than in the forward condition because participants only had to distinguish between four rather than six people.) A Newman-Keuls test showed a significant difference ($\alpha = .05$) between +0.5 and +1 exaggerations (performance was worse for +1 exaggerations), but no other significant differences. An orthogonal contrast was not consistent with a linear relationship, $F(1, 33) = 1.1, p > .1$, the nonlinear residual being significant, $F(2, 33) = 5.3, p < .05$. There was a main effect of actor in this experiment, $F(3, 33) = 6.2, p < .05$, but this was independent of level of exaggeration, $p > .1$. In summary, the very different pattern of results found when sequences were played backward is consistent with interpreting the results of the main experiment as being specific to the perception of biological motion.

---

1. The stimuli used are convincing depictions of biological motion although we used only six points and represented only part of the body. To test
Indeed, the widespread use of key frames by animators suggests that they can be seen when sequences are presented simultaneously. In the case of face recognition, it has been suggested that exaggerations are associated with advantages because faces are stored as deviations from the average in a hypothetical “face space” (Valentine, 1991; for a review, see Rhodes, 1996). The representations might themselves be exaggerations; this would directly explain improved performance, as exaggerations would be a closer match than nonexaggerated examples. Alternatively, exaggerations might provide better access than unexaggerated examples to representations that are themselves veridical (Rhodes et al., 1987); this would be possible within a scheme in which both the local density of the representation space and the distance between representations are important (Krumhansl, 1978). The results presented here are consistent with the possibility that movements are encoded relative to an average spatiotemporal template or prototype (Posner & Keele, 1968). However, the effects of exaggeration can be accounted for within other representational schemes, for example, an exemplar-based account (Medin & Schaffer, 1978; Nosofsky, 1986). It is beyond the scope of this article to determine the nature of the representation involved, but the results do show that temporal information is diagnostic and must be accessible from whatever representation is used (cf. Schyns, 1998). The current results also extend the face-caricature effect to a different domain and suggest that exaggeration may reflect general principles of how information is encoded for within-class discriminations.

For the most part, participants did not report being aware that the timing was being varied. With sequential presentation, as used in this experiment, differences are not, in general, readily apparent, although they can be seen when sequences are presented simultaneously. In the cases in which there was an obvious difference, such as when the glass was picked up particularly fast, this was reliably associated with a particular person and so provided a cue to recognition regardless of how natural it looked. Similarly, even when facial caricatures are obviously distorted, they are still clearly recognizable.

Although the movement, task, and stimuli used here were relatively arbitrary, we think that the results reflect general properties of how biological motion is perceived. It should be possible to temporally exaggerate any movement that can be divided into segments. Indeed, the widespread use of key frames by animators suggests that this would normally be the case. Exaggerating the durations of these segments would be expected to enhance performance on tasks requiring discrimination between the motions involved, at least when temporal or spatiotemporal properties are diagnostic. Our recognition task is just a particular case of a within-category discrimination, and we would expect that the physical and psychological dimensions varied here would be important for other tasks involving the categorization of biological motion.

In conclusion, temporal properties are important to how biological motion is perceived, and exaggerating differences appears to facilitate recognition on the basis of biological motion. The results generalize the effect of exaggeration to another domain and are consistent with exaggeration reflecting general properties of the way information for within-class categorizations is encoded or accessed.

Acknowledgments—We would like to thank the actors and participants who took part in the experiment. We also thank Pascal Mamassian, Philippe Schyns, Donald Morrison, and the two reviewers for helpful and insightful comments on an earlier draft of the article, and Alan Johnston for related discussions. The work was supported by Engineering and Physical Science Research Council Grant GR/M36052 to Frank Pollick.

REFERENCES

Temporal Exaggeration


(RECEIVED 4/28/99; ACCEPTED 8/16/99)