# Toward Collaboration Sensing: Multimodal Detection of the Chameleon Effect in Collaborative Learning Settings

Bertrand Schneider
Stanford University
schneibe@stanford.edu

## ABSTRACT
In this paper, I describe part of my doctoral dissertation in which I have attempted to automatically detect a phenomenon known as the *chameleon effect* in collaborative learning settings. The chameleon effect refers to non-conscious mimicry of other comportments (e.g. postures, mannerisms, facial expressions), such that one's behavior passively and unintentionally changes to match a partner's behaviors. As described below, social mimicry is associated with more productive collaborations and potentially higher learning gains in classroom settings. I describe several studies where I was able to show that visual synchronization (i.e. joint attention), and verbal synchronization (i.e. discourse coherence) were associated with higher learning gains and better collaboration in groups of students, while body synchronization and grammatical mimicry did not predict any of those outcomes. I conclude by discussing implications for educational data mining and describe future work using additional measures such as voice synchronization (e.g. variations in pitch and volume) and arousal synchronization (i.e. variations in heart beat rhythms).

## Keywords
Learning Analytics; Collaborative Learning; Mimicry effect.

## 1. INTRODUCTION
Over the past decades, collaborative learning has been seen as one of the most promising approaches for fostering deep conceptual understanding of complex science concepts. However, even though educational researchers and psychologists have constructed a rich corpus of studies showing the promises of socio-constructivism, much remains to be leaned about effective collaboration among students. As Dillenbourg puts it [3], collaboration in itself is neither good nor bad; there exists conditions that can support productive interactions between students and it's the goal of researchers to discover them. Moreover, he suggests that studies should focus more on process variables rather than learning outcomes: "empirical studies have more recently started to focus less on establishing parameters for effective collaboration and more on trying to understand the role which such variables play in mediating interaction. […] we argue that this shift to a more process-oriented account requires new tools for analyzing and modeling interactions". This is precisely the approach that I am taking in this paper: I use new technologies such as sensors (e.g. eye-trackers, Kinects) combined with data mining algorithms to discover new patterns in collaborative learning settings. More specifically, I used network analysis, natural language processing, supervised and unsupervised machine learning algorithms to make sense of transcripts, eye tracking, and gesture data. My approach is theory driven in the sense that I take advantage of concepts in psychology, ethology and the learning sciences to drive my analyses. One concept that I am closely looking at is the *chameleon effect*.

## 2. THE CHAMELEON EFFECT
The chameleon effect is defined as "the nonconscious mimicry of the postures, mannerisms, facial expressions and other behaviors of one's interaction partners, such that one's behavior passively and unintentionally changes to match that of others in one's current social environment." [1]. The main hypothesis behind my work is that a high level of mimicry in a small group of students is associated with more productive interactions (not only in terms of learning gains, but also in terms of students' quality of collaboration). I do not postulate a causal link between those two variables, even though I showed in one experiment [10] that it is possible to create interventions supporting collaborative learning groups by increasing synchronization between students. On a more theoretical level, previous literature has shown that the chameleon effect is indeed associated with more satisfying and productive interactions. In the next sections, I will briefly summarize the literature suggesting that enhanced levels of coordination support students' learning for each type of synchronization (visual, verbal and postural). I will then present my results and conclude by mentioning implications for designing learning environments.

### 2.1 Visual Coordination
The first example of synchronization is *visual coordination*. Historically, there is a plethora of work (summarized in [10]) showing that joint attention plays a crucial role in any kind of social interaction: From babies learning from their caregivers to parents educating their children, students learning from teachers, students collaborating on a project or for any group of adults working toward a common goal, joint attention is a fundamental mechanism for establishing common ground between individuals.

In my experiment, 21 dyads (N=42) remotely worked on a set of contrasting cases; students had to discover how the human brain processes visual information. The experiment had four distinct steps: first, students were welcomed and assigned to two different rooms. They then took a pre-test measuring their existing knowledge on the topic taught (step 1). In the second step, they collaborated via a microphone when working on the contrasting cases.

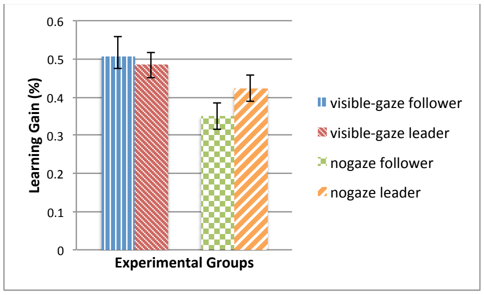


**Figure 1: Results of the experiment conducted in [10].**

In one condition, members of the dyads saw the gaze of their partner on the screen; in a control group, they did not have access to this information. They spent 15 minutes trying to predict how different lesions would affect the visual field of a human brain. In the third step, they then read a text for another 15 minutes on the same topic describing how the visual pathways of the brain work. Finally, they individually took a learning test to assess their understanding of the topic (step 4).

Results indicate that this intervention helped students achieve a higher quality of collaboration, as measured by [10] ($F(1,10) = 24.68$, $p < 0.001$) and a higher learning gain ($F(1,40) = 7.81$, $p < 0.01$). Additionally there was an interaction effect between two factors (experimental conditions and a follower or a leader) on the total learning score: $F(1,38) = 5.29$, $p < 0.05$. Followers learnt significantly more when they could see the gaze of the leader on the screen. They learnt less when they could not (for more detail, see [10]). Interestingly, participants in the "visible-gaze" condition achieved joint attention more often than the participants in the "no-gaze" condition: $F(1,30) = 22.45$, $p < 0.001$. More importantly for the context of this paper, *the percentage of joint attention was one of the only measures correlated with a positive learning gain*: $r = 0.39$, $p < 0.05$. That is, visual coordination was our best measure for predicting students' learning.

I then used network analysis techniques to further exploit this dataset. To construct graphs from gaze data, I divided the screen students had to study into 44 different areas (for more details, see [7]). In this approach, the node size in the dyad graphs is proportional to the number of times dyad members looked at the respective screen area at the same time. Edges are created between nodes when we observe saccades between the corresponding screen regions. The weight of an edge is proportional to the number of saccades between the corresponding screen end-points. Small graphs with few nodes are characteristic of poor collaboration (Fig. 2, left side), and large graphs with highly connected nodes show productive dyads (Fig. 2, right side).

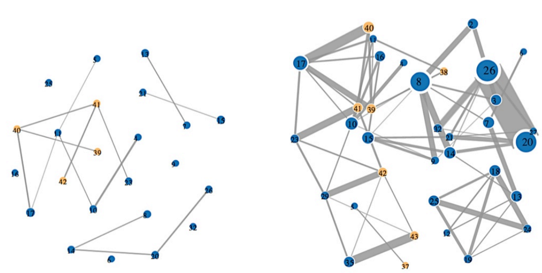

**Figure 2: Graphs based on dyads' data. The size of each node reflects the number of moments of joint attention members of the group shared on one area of the screen.**

Based on this new dataset, we computed various network metrics. I found that in the visible-gaze condition, there were significantly more nodes ($F(1,30) = 8.57$, $p = 0.06$), with bigger average size ($F(1,30) = 22.15$, $p < 0.001$), more edges ($F(1,30) = 5.63$, $p = 0.024$), and more reciprocated edges ($F(1,30) = 7.31$, $p = 0.011$). Those results indicate that *we can potentially separate our two experimental conditions solely based on network characteristics*. In [7], I also show that various network metrics correlate with different aspects of a good collaboration (e.g. the number of nodes (and edges) in the graph were associated with a better ability to reach consensus; betweenness centrality was correlated with the ability of students to sustain mutual understanding; and so on).

In summary, this first study shows that visual coordination is indicative of productive interactions in small collaborative learning groups. I will now turn to the second example, verbal coordination among students.

## 2.2 Verbal Coordination

Danescu [2] mentions how verbal coordination has been shown to enhance communication in organizational contexts, psychotherapy, care of the mentally disabled, and police-community interactions. Thus, there is some evidence showing that verbal mimicry leads to productive interactions. Moreover, I can further divide this concept in two different categories: what Danescu calls *convergence* (i.e. superficial coordination, such as grammatical resemblance) and what other researchers call *coherence* [3] (i.e. deep coordination, such as repeating ideas being expressed by a partner).

Concretely, Danescu used 9 categories from the LIWC corpus (Linguistic Inquiry and Word Counts - http://www.liwc.net/) to compute converge measures. Those categories are: articles, auxiliary verbs, conjunctions, high-frequency adverbs, impersonal pronouns, negations, personal pronouns, prepositions, and quantifiers. The way convergence is computed is relatively trivial:

$$P(b^t_{\hookrightarrow a} = 1 | a^t = 1) - P(b^t_{\hookrightarrow a} = 1).$$

The first expression is the conditional probability of seeing word type *t* expressed by person *b* in answer to person *a*, given that *a* used this word type in the previous utterance. The second expression is just the probability of seeing a particular word type in the entire corpus. Subtracting the second expression from the first one gives us a measure of *convergence*.

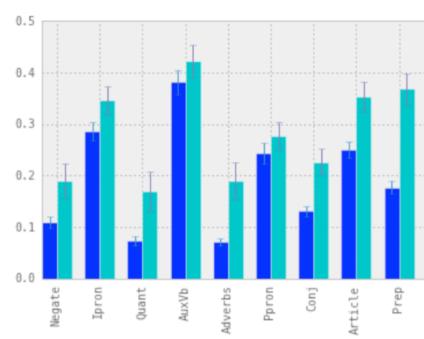


**Figure 3: A replication of Danescu's results on my dataset. Errors bars show standard errors. Non-overlapping error bars show statistically significant differences. Light blue bars show the conditional probability of using a particular word type, given that an interlocutor used it in the previous utterance. Dark blue bars show the probability of using a particular word type in the entire corpus.**

Reusing the dataset from [10], I was able to show that students were indeed mimicking their partners' grammatical structure (Fig. 3). However this measure was not correlated with students' learning gains or quality of collaboration. I then computed the *coherence* of students' discussion (for more details, see [11]): by segmenting the transcripts and computing document similarity measures between those sequential segments (i.e. tf-idf, followed cosine similarity measures), I was able to compute the extent to which students were reusing ideas cited earlier in their discussion. I found that students in the "visible-gaze" condition were significantly more coherent than students in the "no-gaze" condition and that this measure was positively correlated with students' learning gain: $r(19) = 0.540$, $p = 0.011$.

In summary, those results suggest that not any kind of synchronization is indicative of productive patterns of collaboration. I found that *coherence* was associated with higher learning gains, but *convergence* was not.

## 2.3 Postural Coordination

In previous research, I was able to show that joint attention was beneficial to establishing a common ground, which in turn positively influenced how much students learned during an activity [10]. Other lines of research (in ethology as well as in human psychology [1]) suggest that body synchronization is also associated with more productive collaborations. I was inspired by those results and decided to compute a metric for gestures synchronization using the Kinect data. The dataset comes from a study conducted with a tangible interface, where students had to reconstruct the human hearing system [8].

My approach was to first take pairs of data points (one from each student) and computes the distance between them. Distance was calculated by taking the absolute value of the difference between the joint angles of each participant. Those differences were then averaged for each time point. I created graphs with time series of those data points as well as an overall measure of body synchronization. Statistical analyses did not reveal any significant correlation between body synchronization and learning gains: $r(16) = 0.189$, $p = 0.453$. I thus conducted a second attempt that was inspired from the literature in eye-tracking studies: it usually takes +/- 2 seconds for participants in a collaborative situation to adjust their gaze to their partner's behavior. It is possible that body language obeys the same rules. Thus, I repeated the procedure above, but this time, for each data point we looked at the minimum distance in their partner body posture +/- 2 seconds. The correlation with students' learning gains did not reach significance: $r(16) = 0.184$, $p = 0.466$. It suggests that even though gaze synchronization is a strong predictor for students' quality of collaboration and learning, body synchronization does not hold the same properties, at least in the context of this experiment. Successful students were *not* more likely to coordinate their action based on their partner's behavior.

# 3. DISCUSSION

The work described above shows a first step in computing multimodal metrics of the *chameleon effect*. I showed how visual coordination and verbal coordination were associated with higher learning gains. I also found that grammatical coordination and body synchronization was not significantly correlated with students' quality of collaboration or learning gains. This means that the chameleon effect is not universal: at least in educational settings, it varies in its form and intensity according to different modalities.

Future work should focus on alternative measures of the chameleon effect (e.g. voice features, heart beat rhythms) and assess whether other kinds of synchronization are associated with positive learning outcomes. Future work should also explore the approach described in [7] to a greater extent: building network or probabilistic models on top of large datasets is likely to lead to additional insights in terms of students learning processes. Implications of this work are manifold. We can imagine feeding those features into a machine learning algorithms to predict students' quality of collaboration; this prediction can then be used by a teacher of by a learning environment to propose various scaffolds supporting students' learning. Finally, those metrics can potentially lead to a greater understanding of human social interactions by isolating where and when the chameleon effect actually applies. This understanding can lead to the development of new feedback loops, such as the one described in [10] (i.e. the gaze-awareness tool used by students).

# 4. REFERENCES


[1] Chartrand, T.L. and Bargh, J.A. The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology 76*, 6 (1999), 893–910.

[2] Danescu-Niculescu-Mizil, Cristian, and Lillian Lee. "Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs." *arXiv preprint arXiv:1106.3077* (2011).

[3] Dillenbourg, P. (1999). What do you mean by "collaborative learning"? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-16). Amsterdam, NL: Pergamon, Elsevier Science.

[4] Graesser,A.C.,McNamara,D.S.,Louwerse,M.M.,and Cai, Z. "Coh-Metrix: Analysis of text on cohesion and language." *Behavior Research Methods, Instruments, & Computers 36*, 2 (2004), 193–202.

[5] Meier, Anne, Hans Spada, and Nikol Rummel. "A rating scheme for assessing the quality of computer-supported collaboration processes." *International Journal of Computer-Supported Collaborative Learning, 2*, 1 (2007), 63-86.

[6] Pennebaker, J W., M. E. Francis, and R. J. Booth. "Linguistic inquiry and word count: LIWC 2001." *Mahway, NJ: Lawrence Erlbaum Associates,* 2001, 71.

[7] Schneider, B., Abu-El-Haija, S., Reesman, J., & Pea, R. (2013). Toward Collaboration Sensing: Applying Network Analysis Techniques to Collaborative Eye-tracking Data. *ACM International Conference on Learning Analytics*, LAK '13 (pp. 107-111). Leuven, Belgium: ACM.

[8] Schneider, B., & Blikstein, P. (submitted). Computational Techniques for Detecting Productive Interaction Around an Interactive Tabletop. *EDM2014.*

[9] Schneider, B., & Pea, R. (in preparation). Toward Collaboration Sensing. *International Journal of Computer-Supported Collaborative learning.*

[10] Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 375-397.

[11] Schneider, B., & Pea, R. (submitted). Exploring the Effect of Mutual Gaze Perception on Students' Verbal Coordination. *EDM2014.*