

# Make gestures to learn: Reproducing gestures improves the learning of anatomical knowledge more than just seeing gestures

Mélaine Cherdieu<sup>1</sup>, Olivier Palombi<sup>2,3</sup>, Silvain Gerber<sup>1</sup>, Jocelyne Troccaz<sup>4</sup> and Amélie Rochet-Capellan<sup>1\*</sup>

<sup>1</sup> Univ. Grenoble Alpes, CNRS, GIPSA-lab, F-38000 Grenoble, France

<sup>2</sup> Univ. Grenoble Alpes, University Hospital, Anatomy Laboratory, LADAF, F-38000 Grenoble, France

<sup>3</sup> Univ. Grenoble Alpes, CNRS, LJK, F-38000 Grenoble, France

<sup>4</sup> Univ. Grenoble Alpes, CNRS, TIMC-IMAG, F-38000 Grenoble, France

## \* Correspondence:

Amélie Rochet-Capellan

[amelie.rochet-capellan@gipsa-lab.grenoble-inp.fr](mailto:amelie.rochet-capellan@gipsa-lab.grenoble-inp.fr)

**Number of words (introduction to references): 7882; Number of figures: 6**

**Keywords: Manual gesture; Learning and memory; Embodied cognition and learning; Anatomy; Long-term memory**

## Abstract

Manual gestures can facilitate problem solving but also language or conceptual learning. Both seeing and making the gestures seems to be beneficial to cognitive tasks and learning. However, the stronger activation of the motor system in the second case should provide supplementary cues to consolidate and re-enact the mental traces created during learning. We tested this hypothesis in the context of anatomy learning by novice adults. Anatomy is a challenging topic to learn and is of specific interest for research on embodied learning, as the learning content can be directly linked to learners' body. Two groups of participants were asked to look at a video lecture on the forearm anatomy. The video included a model making gestures related to the content of the lecture. Both groups see the gestures but only one also imitate the model. Tests of knowledge were run just after learning and few days later. The results revealed that doing gestures improves the recall of structures names and their localization on a diagram. This effect was however significant only in long-term assessments. This suggests that: (1) the integration of motor actions and knowledge may require sleep; (2) a specific activation of the motor system during learning can improve the consolidation and/or the retrieval of memories.

## 1 Introduction

The movements we make do not only act on the outer world but they also shape our cognitive abilities. A broad range of neurological and behavioral evidences suggest that the motor system is involved in perception, decision making, lexical representation, memory and emotional processes (Barsalou, 2008; Casasanto & Dijkstra, 2010; Glenberg & Kaschak, 2002; Tyler et al., 2003; Wexler, Kosslyn, & Berthoz, 1998). Among the various body actions shown to influence our cognitive abilities, research paid particular attention to manual gestures, due to their multi-faceted representative function and their special relationship to language (Goldin-Meadow & Alibali, 2013; McNeill, 1992). Manual gestures are commonly and spontaneously used by speakers, learners and teachers. They were shown to support language learning on receptive and expressive sides as well as conceptual learning and problem solving (Kontra, Goldin-Meadow, & Beilock, 2012; Novack & Goldin-Meadow, 2015). Experimental results also indicate that when children or adults gesture during learning they outperform participants who do not. Embodied and active learning is therefore a major focus, at the crossroad of cognitive and education sciences but mainly investigated with children so far, in regards to educational issues (Kelly, Manning, & Rodak, 2008; Kontra et al., 2012; Novack & Goldin-Meadow, 2015). The role of gesturing in learning might depend on the topic of interest and population. In this paper, we evaluate the beneficial aspect of gesturing in anatomy learning in novice adults. Anatomy is a challenging topic for students (Sugand, Abrahams, & Khurana, 2010; Zumwalt et al., 2010), involving visuo-spatial skills, mental imagery, as well as lexical memories. Making gestures to learn anatomy is also a specific case of embodied active learning, in which explicit knowledge on the body is developed throughout the implication of the body. Our experimental work compares anatomy learning of the forearm in controlled situations when seeing vs. seeing and imitating gestures related to the educational content.

Sensory-motor experiences, body movements, are a growing focus of researches in various fields of cognitive sciences. These researches have been developed in conjunction with the theoretical framework of “embodied cognition” (cf. Barsalou, 2008; Borghi & Pecher, 2011; Clark, 1999; Varela, Thompson, & Rosch, 1993; Wilson, 2002). Nowadays, researches on the implication of sensory-motor experiences in cognitive functions are clearly connected to and/or motivated by embodied cognition (Hostetter & Alibali, 2008; Ionescu & Vasc, 2014). This is also the case of the present paper. According to the embodied cognition framework, mental functions are grounded in sensory-motor processes linked to specific situations (“grounded cognition”, see Barsalou, 2008). Therefore, cognitive representations are not amodal but rather “grounded” in the sensory-motor processes involved in their construction. In knowledge acquisition, as in all experiences, “the brain captures states across the modalities and integrates them with a multimodal representation stored in memory” (Barsalou, 2008, p.618). In this framework, cognitive processes rely on simulation mechanisms that could be defined as “the reenactment of perceptual, motor, and introspective states acquired during experience with the world, body, and mind” (Barsalou, 2008, p.619). As specific cases of sensory-motor experiences, and due to their powerful representational and deictic functions, manual gestures are the core interest of a variety of studies focused for instance on the role of gestures in language, conceptual learning, problem solving or spatial representations. The positive effect of gestures in these cognitive activities has been shown when seeing the gestures made by another person (a teacher, a discussion partner etc.) and when making the gestures.

A major human activity that clearly involves gestures is communication. In a daily basis, people are spontaneously gesturing in order to communicate. These gestures can mark the rhythm of the discourse, strengthen the message or provide additional information to the listener. In a recent meta-analysis, Hostetter (2011) highlighted the importance of gestures in receptive communication. Seeing

speaker's gestures helps listeners to understand the spoken content of the message. This effect is more important when speaker's gestures provide supplemental information rather than redundant information to the spoken message, and it is especially relevant for children. These results suggest that gestures could represent "a second channel that makes successful comprehension more likely" (Goldin-Meadow & Alibali, 2013, p.261). Gestures are thus seen as a facet of human language (Kelly et al., 2008; McNeill, 1981). The effect of gestures on language comprehension however depends on the information contained in the gestures (Goldin-Meadow & Alibali, 2013), and on the complexity of the spoken message (McNeil, Alibali, & Evans, 2000). Besides, the increase of spoken message by gestures has beneficial effect in education. During class, teachers spontaneously gesture when they have to explain mathematical concepts to children (Goldin-Meadow, Kim, & Singer, 1999) and several studies have shown that seeing the teacher's gestures during learning is beneficial to the learner. When children see gestures while they learn to solve problems in mathematics, they improve more between pre-test and post-test than when they received oral instructions only (Cook, Duffy, & Fenn, 2013). This has also been observed for concepts learning such as the concept of conservation (comparing the same amount of water in two different glasses, Church, Ayman-Nolley, & Mahootian, 2004; Ping & Goldin-Meadow, 2008) or symmetry (Valenzeno, Alibali, & Klatzky, 2003). Interestingly, Singer & Goldin-Meadow (2005) found that seeing gestures during a lesson of mathematics improves performance 24 hours after the lesson and tends to facilitate transfer to other problems solving. Teachers' gestures can help children understand and integrate instruction, which can explain the positive effect on long-term performances. Here again, gestures should convey supplemental information to verbal information to be helpful: when children learn mathematics, they only improved when the teachers' gestures bring a different way to approach the problem than the teacher's spoken message. These results highlight the crucial role of gestures in language comprehension and in different types of conceptual learning. However, if seeing gestures is important, doing gestures also has a major impact in cognitive activities.

Gestures appear very early in the development of expressive communication, even before the first words. For example, children can point objects before they can name them. Not only gestures precede oral language, they seem related to it over the course of development: the early use of gestures in communication can predict some aspects of speech acquisition. In a longitudinal study, Iverson & Goldin-Meadow (2005) showed that the spoken lexical items developed by children around 24 months are tied to the gestures they have made during their early communicative activities, around 10 months. Early gestures do not only predict the lexical items, they also predict the vocabulary range (Iverson & Goldin-Meadow, 2005). Thereby, gestures made by 14 months children predict their vocabulary size at 42 months (Rowe & Goldin-Meadow, 2009). Children also use gestures to organize their thoughts and solve problems. For instance, in Church & Goldin-Meadow (1986), children had to explain their understanding of the concept of conservation while judging the same amount of water poured into two glasses with different shapes. When they explained their reasoning to solve the problem, some of them spontaneously produced gestures matching their reasoning while others produced gestures mismatching their reasoning. The latter were more likely to understand their mistakes, and better improve their performance than the former. The role of making gestures in the organization of thought in children has also been shown for mathematical problem solving (Alibali & Nathan, 2012; Goldin-Meadow & Beilock, 2010) and moral reasoning (Beaudoin-Ryan & Goldin-Meadow, 2014). Besides supporting language production and reasoning, gestures are more generally involved in knowledge access. In a recent study, Gunderson et al. (2015) reported that when children gestured during number-knowledge tasks (tasks involving the evaluation of set sizes), they could have access to knowledge not included in their spoken message, that is to say "numbers above their knower-level" (Gunderson et al., 2015). In the same way, tracing during geometry learning enhances the children's knowledge, who then perform better in evaluation tests

(Hu, Ginns, & Bobis, 2015). Gestures therefore appear as a supplemental dimension to learning, helping children who gesture to outperform children who don't gesture. This influence of gestures on thought, problem solving and knowledge access is yet not restricted to children. Rauscher, Krauss, & Chen (1996) have shown that gesturing while speaking could facilitate access to the mental lexicon in spontaneous speech produced by undergraduate students (Rauscher, Krauss, & Chen, 1996). The positive influence of gestures was also assessed in learning new words, throughout second language learning (McCafferty, 2004): adults can learn a second language easier if they gesture during learning (Gullberg, 2006; McCafferty, 2004). As observed in children, gestures are also an important part of the organization of thought and reasoning in adults. In a recent study (Trofatter, Kontra, Beilock, & Goldin-Meadow, 2014), a group of students had to solve the Tower of Hanoi puzzle. This puzzle is used in neuropsychology to assess planning and reasoning. The students had to solve the problem once and then had to explain their reasoning. The students who made gestures during their explanation showed a greater improvement when they solved the Tower of Hanoi again than those who did not make gestures. Gestures are quite important in learning, reasoning and language and should surely be taken into account in education for children as well as for adults (Ionescu & Vasc, 2014). These previous studies have recently inspired researches about the role of gestures in anatomy learning.

Anatomy learning is particularly challenging for medical and paramedical students. Functional anatomy requires recalling the names, locations and functions of the different structures of the human body. Various studies investigate new teaching methods inspired by embodied cognition in order to facilitate anatomy learning. Therefore painting anatomical structures on the learner's body (McMenamin, 2008) or drawings them on t-shirts (Skinder-Meredith, 2010) can have a positive effect on student's motivation and consequently possibly enhance performances (for a review see Sugand et al., 2010). However, most of these studies did not objectively investigate the effect of these methods on anatomy tests scores. The potential role of gestures in anatomy learning is particularly interesting because it is about learning knowledge on the body throughout the body. Intuitively we can think that gestures will facilitate the memorization of anatomical structures and better understand their implication in body functions. This point has been developed in some studies using gestures in different ways. Oh, Won, Kim, & Jang (2011) described what they called "digit anatomy": students had to reproduce several fingers configuration to learn the anatomy of the heart. In this case, the fingers represented vessels, bones and muscles. A subjective evaluation of the technique indicated that students found the method helpful to memorize the structures and to understand the functions. Pointing and tracing gestures were also tested as supports of anatomy learning in Macken & Ginns (2014). During the learning of the structures and functions of the human heart, some students were told to do pointing and tracing gestures on diagrams, while others were told not to use their hands (they had to sit on their hands to control they would not make any gestures). Students who gestured during learning had better performance than those who did not, both for the terminology ("knowledge of specific definitions and labels") and comprehension (understanding of the structures and functions of the heart) tests (Macken & Ginns, 2014; p 598). This result suggests that pointing and tracing gestures can improve the learning of heart anatomy. These previous studies investigated the role of different types of gestures in anatomy learning of internal structures, in particular the heart. Others authors introduced the possibility to make movements to learn information about the structures and the functions involved in these particular movements. Therefore, Dickson and Stephens (2015) investigated if making facial expressions could improve the learning of cranial nerves involved in these expressions. Students were taught a lesson about these cranial nerves without doing any gesture (Didactic lecture) and then attended to the same lecture when asked to do facial expressions illustrating the functions of the different nerves in facial movements (Miming lecture). A pre-test and a post-test of knowledge were administered before and

after each lecture. When comparing the results at pre-tests and post-tests the authors observed a greater improvement for the Miming condition as compared with the Didactic condition. It is however unclear if this improvement is related to the implication of movements or if it is only related to the fact that the lesson was received twice. Movements were also done collectively without any control of students' behavior. The study was acknowledged as a pilot study.

Based on these previous researches in cognitive sciences and education, our aim was to provide further evidences for the role of making gestures in the acquisition of new knowledge by adult population. To this end, we investigated the effect of arm-hand gestures on the learning of structural and functional anatomy of the forearm. Our main hypothesis was that, as compared with just seeing the gestures made by someone else, making gestures during learning could improve the memorization of the structures names (gestures as part of lexical representation and/or as cue to lexical access), structures location (gestures as spatial cues) and/or the understanding of functions (gestures as support for mental imagery and conceptual learning). Indeed, a stronger activation of the motor system in the former case may create additional components to the memory traces and/or contribute to shape abilities to mentally simulate the function (Barsalou, 2008). The effect of making gestures was assessed by comparing two groups of participants (seeing vs. making gestures) both in terms of performance and subjective evaluation. All the evaluations were done after the learning session and a subpart of them few days after, as previous work suggest that gestures "make learning last" (Cook et al., 2008). Gesture behaviors were also monitored during the learning and test phases and the quantity of gestures made was related to the scores in the different evaluations.

## 2 Material and Methods

### 2.1 Participants

Forty-five adult volunteers (aged 18 to 44, 27 females) gave written informed consent to participate in the study. They were recruited by advertisements and a dedicated volunteers' database at Grenoble-Alpes University. Participants were informed of the inclusion criteria: French native speaking, no neurological or psychiatric disorders, normal or corrected to normal sight and hearing, novice knowledge of anatomy (to avoid ceiling effects and too important between-subjects differences in performance). All of them received a 15 € shopping gift card for their participation. The study was approved by the CERNI (Ethic Committee for Non Interventional Research) associated to the Grenoble-Alpes University.

### 2.2 Distribution of participants into two comparable groups

Our experimental design required two groups of participants: a Gesture group and a Control group in which participants do or do not make gestures during learning respectively. The distribution of the participants into these two groups was done such as to balance various skills that could influence performances in anatomy learning (Berney, Bétrancourt, Molinari, & Hoyek, 2015; Hoyek et al., 2009). The following skills were evaluated during a pre-test phase:

- Visuo-spatial abilities, measured using a redrawn version of the Mental Rotation Test (MRT) from Vanderberg and Kuse (1878). We selected the first 12 items of the original test so that this pre-test was not too long for the participants (Peters et al., 1995).
- Mental imagery abilities, evaluated using a French translated version of the Movement Imagery Questionnaire-Revised (MIQ-R) from Hall and Martin (1997) introduced by Lorant and

colleagues (Lorant & Nicolas, 2004). This test consisted in eight questions in which participants had to evaluate their ability to mentally represent themselves doing a movement.

- Ability to memorize complex unknown words, assessed using a list of 8 pseudo-words with two or three syllables that we generated from the Lexique 2 database (New, Pallier, Ferrand, & Matos, 2001). The pseudo-words were selected to have different orthographic organization.

Participants were assigned to one of the two groups using profile pair matching according to their scores in pre-tests. As far as possible, the age, gender and educational level were also taken into account. The final groups included 21 participants each, aged  $25.6 \pm 5.62$  (mean  $\pm$  standard deviation) years old in the Control group (13 females) and aged  $27.7 \pm 6.86$  years old in the Gesture group (12 females).

## 2.3 Learning and testing material

The educational material was a simplified lecture about the structural and functional anatomy involved in supination and pronation movements of the forearm, and associated tests of knowledge. As illustrated in Table 1, supination refers to an external rotation of the forearm and pronation to an internal rotation of the forearm (palm of the hand upwards and downwards at the end of the movement, respectively). This topic was chosen because the lesson could easily be simplified and participants' own forearms being within sight, they could be directly linked to the learning content. The material was elaborated in collaboration with a professor of anatomy and take into account both pedagogical and experimental objectives.

### 2.3.1 Learning material

For the learning phase, a video was designed from educational slides that displayed the anatomical diagrams representing different parts of the forearm bones, joints and ligaments. A soundtrack with explanations from the professor was synchronized with the slides. On the top-right of each slide, a video of the experimenter (the "model") indicated the gestures to make or to observe. This video was always visible with the model staying still during audio explanations. An overview of the lesson is provided in Table 1. The lesson was 10 min long and included 8 slides and 10 gestures.

\*\*\* table 1 \*\*\*

The gestures were inspired from those used by the professor in his real courses. They were mainly pointing-like and iconic-mime gestures (Abner, Cooperrider, & Goldin-Meadow, 2015; Kelly et al., 2008). They were designed such as to convey complementary information that should help understanding: the function (e.g. step 1 in Table 1); the location of a structure in the forearm (e.g. step 2 in Table 1); the structure itself (e.g. step 3 in Table 1); or the implication of the structure in the function (e.g. Step 3 to 5 in Table 1).

At the end of each slide, the video showing the gestures was masked with a green square indicating to the participant the gesture will be played again and she would have to redo it (Gesture group) or to watch it again while staying still (Control group). The video was the same for the two groups except for the instructions written in the green screen.

### 2.3.2 Testing material

Learning of the lesson was assessed using four paper and pencil tests chosen to assess different levels of memory and knowledge:

- *Recall*: participants write down on a white sheet, as many as possible of the 16 structures names saw during the learning phase (cf. the list of names on Figure 1).
- *Diagram*: participants write down the names of structures at the correct location indicated on the diagrams (cf. diagrams on Figure 1, without the numbers).
- *Diagram\_Name*: participants localize the names of the structures on the diagrams (cf. Figure 1). The diagrams were provided together with the list of the structures names in random order. Participants had to write the number corresponding to the correct location in front of the name of the appropriate structure.
- *Multiple Choice Questions (MCQ)*: participants had to choose the correct answer to six questions related to the understanding of functional aspects. For each question, participants circled the answer among four propositions but only one correct.
- *Subjective evaluation*: participants had to evaluate the clarity, motivation, interactivity, engagement of the lecture, and the usefulness of the gestures by evaluating their level of agreement with five assertions, one for each dimension (for instance, “I felt engaged during learning” to assess engagement), on a Likert scale ranging from 1 (I strongly disagree, always the most negative judgment) to 10 (I fully agree, always the most positive judgment).

\*\*\* Figure 1 \*\*\*

## 2.4 Procedure

The experiment was about one-hour long. The participants were informed of the experimental conditions before starting the experiment. They filled a consent form and, gave or not their consent to be videotaped during the learning and test phase. The participants who agreed to be filmed (n=33), signed an image-right form which specified that the recordings will be used for the purpose of the study only. The procedure is summarized on Figure 2 and included the following steps:

\*\*\* Figure 2 \*\*\*

### 2.4.1 Pre-tests phase

The instructions of the two visuo-spatial tests were given and participants complete the Movement Imagery Questionnaire-Revised (MIQ-R) and the Mental Rotation Test (MRT). They then filled a questionnaire to assess their previous knowledge in anatomy and more specifically relative to the content of the lesson (cf. complete the diagram displayed on Figure 1, left). They finally performed the words memory test. In this last test, the 8 pseudo-words were written on a sheet and the experimenter read them along with the participants. They had one minute to memorize the pseudo-words and to retrieve them verbally. The experimenter wrote down the words recalled by the participants. Scores in MIQ-R, MRT and words memory pre-tests were used to assign participants to each group: Gesture vs. Control (see above). After the pre-tests phase, a digital camera was set-up to videotape the participants who agreed to be filmed. For participants who didn't agree, the experimenter took notes of the gestures made for each slide during the learning phase and each test during the evaluation phase.

### 2.4.2 Learning phase

During the learning phase participants were seated at a table, in front of a 22-inch screen. All the participants were instructed to listen and watch carefully the videotaped lesson displayed on the screen and to memorize as much information as possible. They were informed that a video of the experimenter would be played on the top right of the screen. In this video, the experimenter would make gestures that should help them to understand the content of the course. Participants in the



Gesture group were instructed that they would have to reproduce the gestures when the instruction “it’s your turn” will be displayed on a green square, as the model will do the gestures again. Participants in the Control group were told to keep their hands on the table during the whole learning phase. Participants listened to the professor’s explanations with headphones.

### 2.4.3 Evaluation phase

Right after the learning phase, participants completed the four tests of knowledge. They were informed that each test was time pressured (five minutes each). The experimenter provided the tests to the participants in the following order: Recall, Diagram, Diagram\_Name and MCQ, on separated paper sheets and respecting the timing. Finally, participants completed the subjective evaluation. At the end of the first session, the experimenter asked to the participants their agreement to call them back 2 or 3 days after the experiment to do a long-term assessment. When the participants agreed (19/21 participants in each group did), a paper-sheet with the diagrams displayed on Figure 1, was provided to them (without the list of names). The instruction was to not revise anything about the lesson before the call and to wait for further instructions.

### 2.4.4 Long-term assessment

The long-term assessment was performed by phone and lasted about 10 min. Phone call was chosen to reduce the loss of participants and preserve the balance between groups. During the phone call, the experimenter informed the participants that they would have to perform the Recall and Diagram tests again, with time pressure as well. First, she asked the participants to recall as many of the structures studied during the learning as possible (Long-Term Recall). She then asked the participants to visualize the diagrams and to give the name of the structure corresponding to each number (Long-Term Diagram). The experimenter wrote down participants’ answers, including mispronunciations.

## 2.5 Scoring in the different evaluations

### 2.5.1 Tests of knowledge

Scoring was crosschecked by the first and the last author. All the scores were expressed as a proportion of correct answers relative to the total number of items.

- *Recall scoring (short and long term)*: 16 names of structures were expected. Each provided item was scored as 1 if identical to one of the expected name, or a synonym/equivalent name also used by the professor in the video (cf. head of the radius for radial head).

- *Diagram test scoring (short and long term)*: an item was rated 1 if the name of the structure was exact and if it was correctly located on the diagrams.

- *Diagram\_Name scoring*: each structure correctly localized on the diagrams was rated 1.

- *Multiple Choices Questionnaire (MCQ)*: one point was given for each correct answer.

For Recall and Diagram, we also computed the proportion of common letters between the expected and recalled words as it was done in a recent study using a similar assessment test (Skulmowski, Pradel, Kühnert, Brunnett, & Rey, 2016). We chose to use the orthographic distance of Levenshtein, an indicator of spelling proximity (Levenshtein, 1966). This more complex analysis led to the same conclusions as the “all-or-nothing” scoring, and so is not reported here.



## 2.5.2 Subjective evaluations

For the subjective evaluations, a score out of 10 was obtained for each dimension and participant.

## 2.5.3 Gesture behaviors

Gestures made by participants were counted using the ELAN software (version 4.9.2, Lausberg & Sloetjes, 2009), if a videotape was available, or based on the experimenter's observations otherwise. For each participant we obtained a number of gestures at each step of the learning phase and for each test of day 1 (Recall, Diagram, Diagram\_Name and MCQ).

## 2.6 Experimental design, hypotheses and statistics

The main response variables were:

- The proportion of correct responses in the different tests (PROP\_CORRECT, values from 0 to 1);
- The five subjective evaluations of the lecture (QUAL\_EVAL, considered as an ordered variable with values from 1 to 10)

The main factors tested were:

- GROUP (Control group vs. Gesture group), between-subject factor;
- MOMENT (Short-term vs. Long-term), within-subject factor;
- TEST\_TYPE (Recall vs. Recall\_Diagram vs. Diagram vs. Diagram\_Name vs. MCQ) within-subject factor;
- DIMENSION (Clarity, Motivation, Interactivity, Engagement, and Usefulness), for the subjective evaluation only.

The following hypotheses were tested using the version 3.2.2 of the R software (R Core Team, 2015):

*Hypothesis A. Making gestures may improve learning; the effect should last in long-term evaluations:* We were expecting a significant effect of GROUP, with greater performances in the Gesture than Control group on PROP\_CORRECT. This effect could vary depending on TYPE\_TEST and on MOMENT with two main expectations: (1) if making gesture improve the mental simulation of the movement it should be particularly beneficial in the MCQ test that evaluates functional anatomy (significant GROUP\*TYPE\_TEST interaction); (2) if making gesture helps consolidating memories more than just seeing gestures, between groups difference should be greater in the long-term evaluation than short-term one (significant GROUP\*MOMENT interaction). As PROP\_CORRECT ranged between 0 and 1, it was analyzed with a Beta regression, a type of multiple linear regression, using the “glmmADMB” package (version 0.8.3.3, Bolker et al., 2012). A correction was applied to the data to get values in ]0;1[ range, as required by the method (Smithson & Verkuilen, 2006). The contribution of each variable was evaluated using a step-by-step descendant procedure with a backward method: we started the analysis with the full model (containing all the variables and all the interactions) and then step-by-step suppressed the terms that were not critical to the model. This was done by comparing models with and without the variable of interest using likelihood ratio tests (the *anova* function of the package *stats*, 2016.3.3.2 in R). Differences between models were considered significant for  $p < .05$ . PARTICIPANT was included as a random factor. As Diagram\_Name and MCQ were not included for the long-term assessment, it was not possible to

include these two tests in the model evaluating the effect of MOMENT. Therefore, we ran two different models:

- Model\_ST (short-term): effect of GROUP and TYPE\_TEST for short-term data only, all the participants, starting with:

$$PROP\_CORRECT \sim GROUP * TEST\_TYPE + (1|PARTICIPANT)$$

- Model\_LT (short-term and long-term): effect of GROUP, TYPE\_TEST\_LT (Recall and Diagram) and MOMENT and participants with long-term data only (PARTICIPANT\_LT, n=19 in each group):

$$PROP\_CORRECT \sim GROUP * TEST\_TYPE\_LT + (1|PARTICIPANT\_LT)$$

The selected models were then used to run multiple comparisons based on contrast matrixes, using the “multcomp” package (version 1.4-6) and the function “glht” in R (Hothorn, Bretz, & Westfall, 2008). This package applied corrections for multiple comparisons.

*Hypothesis B. Making gestures may improve the subjective evaluation of the lecture.* We were expecting that moving their body during the learning phase could improve the way participants evaluate the clarity, motivation, interactivity, engagement, and gestures usefulness dimensions of the lecture. The effect of making gestures on QUAL\_EVAL according to DIMENSION was evaluated using an ordinal regression (function clmm of the package Ordinal in R 2015.6-28, Christensen, 2015). The principle is similar to the Beta regression with a full model at the beginning and a backward step-by-step method. The full initial model was the following:

$$QUAL\_EVAL \sim GROUP * DIMENSION + (1|PARTICIPANT)$$

*Hypothesis C. The number of gestures made during learning and testing phase may influence the performances.* We finally analyzed the number of gestures made during the different phases and when possible, correlate this number to scores in the different evaluations.

### 3 Results

#### 3.1 Constitution of groups

We first controlled that the two groups were balanced in terms of pre-test knowledge, MRT, MIQ, words memory abilities as well as average age, educational level and number of days between short-term and long-term evaluations. The results are reported in Table 2. Mann-Whitney non-parametric tests show no significant differences between groups (without correction, in all cases:  $W < 217$ ,  $p > .22$ ), suggesting that the two groups are globally comparable on the tested dimensions. This was also the case after removing the four participants that didn't performed the long-term assessment. Moreover, none of the participants evaluated herself as being an expert in functional anatomy and anatomy of the upper limb, only the radius was correctly reported and located on the diagram filled before learning by two participants. All the participants are thus considered to be naïve in regards to the content of the lesson.

\*\*\* Table 2 \*\*\*

#### 3.2 Effect of GROUP and TYPE\_TEST in the short-term session

The average performances in the short-term session are displayed on Figure 3. Regardless of TEST\_TYPE, average scores are always greater in the Gesture than in the Control group. Greater performances are also observed in Diagram\_Name and MCQ as compared with Recall and Diagram. The beta regression analysis (Model\_ST) shows that the contribution of GROUP\*TEST\_TYPE is not significant ( $\chi^2(3)=3.9$ ,  $p=.27$ ), nor the global effect of GROUP ( $\chi^2(1)=1.9$ ,  $p=.17$ ). By contrast, TYPE\_TEST significantly improves the model ( $\chi^2(3)=111.3$ ,  $p<.001$ ), with greater scores in MCQ and Diagram\_Name than in Recall and Diagram (for all paired comparisons between the two groups of tests:  $z>5.9$ ,  $p<.001$ ). It is not surprising that participants performed better in the Diagram\_Name than in the Recall and Diagram tests, as the structures names are provided in Diagram\_Name. Comparison with MCQ is less relevant as the level of difficulty is not standardized between this test and the others. The important point here is that TEST\_TYPE effect doesn't significantly depend on the group.

\*\*\* Figure 3 \*\*\*

### 3.3 Effects of GROUP, TYPE\_TEST\_LT and MOMENT

Figure 4 displays the average results for each GROUP and MOMENT when analyses are limited to TYPE\_TEST\_LT and PARTICIPANT\_LT. Short-term results are consistent with those observed when considering all the participants for the Recall and Diagram tests. Long-term results suggest however an increase of differences between the two groups in the long-term session as well as an improvement of scores in the Gesture group mainly as compared with the short-term scores. The beta regression analysis (model\_LT) shows no significant contribution of the three-levels interaction (GROUP\*TYPE\_TEST\_LT\*MOMENT,  $\chi^2(1)=.01$ ,  $p=.91$ ), nor of TYPE\_TEST\_LT\*MOMENT ( $\chi^2(1)=.44$ ,  $p=.51$ ), nor of GROUP\*TYPE\_TEST\_LT ( $\chi^2(1)=1.0$ ,  $p=.31$ ). By contrast, TYPE\_TEST\_LT significantly contributes to the model, with greater performances in Recall than Diagram test ( $\chi^2(1)=8.7$ ,  $p<.01$ ), as well as the GROUP\*MOMENT interaction ( $\chi^2(1)=8.3$ ,  $p<.01$ ). Contrast comparisons were then performed on this last interaction. They show no significant difference between groups for short-term assessments ( $z=1.22$ ;  $p=.29$ ), but significant difference for long-term assessments ( $z=2.36$ ;  $p<.03$ ). Supplementary analysis suggest that the improvement at long-term was significant for the Gesture group ( $z=3.26$ ,  $p<.01$ ) but not the Control group ( $z=1.38$ ,  $p=.28$ ).

\*\*\* Figure 4 \*\*\*

### 3.4 Subjective evaluation

Distributions of subjective evaluation scores (QUAL\_EVAL) are displayed on Figure 5. QUAL\_EVAL is distributed in an equivalent way in the two groups for the clarity of the lecture and the usefulness of the gestures. For the other dimensions (engagement and motivation in the lecture, interactivity of the lecture), participants in the Gesture group seem to provide higher scores than in the Control group. The ordinal regression shows no significant contribution of GROUP\*DIMENSION interaction (Likelihood ratio statistic:  $LR.stat(4)=6.7$ ,  $p=.15$ ), nor of GROUP ( $LR.stat(1)=1.12$ ,  $p=.29$ ). By contrast, QUAL\_EVAL significantly depends on DIMENSION ( $LR.stat(4)=34.25$ ,  $p<.001$ ), with in general, more positive evaluation of Clarity and Usefulness than Motivation, Interactivity and Engagement. Note that when dimensions are compared independently using Mann-Whitney tests, Engagement is significantly better evaluated by the Gesture than the Control group ( $W=135.5$ ,  $p=.03$ ), while there is no significant between-groups differences for the other dimensions ( $W>171.5$ ,  $p>.2$ ). This supplementary analysis was run for comparison purposes

with previous works suggesting that students feel more involved in the lecture when making gesture than not, based on mean comparisons (Oh, Won, Kim, & Jang, 2011).

\*\*\* Figure 5 \*\*\*

### 3.5 Gestures analysis

#### 3.5.1 Gestures made during learning

As described in the methods section (cf. Table 1), the model made 10 gestures during the learning phase. We analyze to which extent the participants reproduced these gestures. The majority of the participants in the Control group stayed still during learning but few of them (N=4) made 1 or 2 gestures (Figure 6, top left), at variable steps of the lesson. We still included them in the control group, as these gestures were occasional and not systematic. In the Gesture group, all the participants made the 10 expected gestures at least, but most of them repeated the gestures several time, with more than half of the participants making at least 21 gestures in total (Figure 6, top right). As an illustration of specific behaviors, the participants G-02 and G-04 made each gesture once, while G-21 repeated each gesture several times, up to 50 gestures in total.

\*\*\* Figure 6 \*\*\*

#### 3.5.2 Gestures made during the evaluation phase

As displayed on Figure 6, second row, during the evaluation phase, participants' behaviors were variable and depended on the test. Globally, the participants poorly gestured when completing Recall (number of participants making at least one gesture: n=4 both in Control and Gesture group), Diagram (n=1 in Control, 0 in Gesture) and Diagram\_Name (n=1 in Control, n=4 in Control). Outlier behaviors are also observed. For example, C-03 and G-10 made 22 and 12 gestures when completing Recall and Diagram\_Name respectively. More gestures were made during MCQ, with about half of the participants in each group (n=11 in Control and n=12 in Gesture) who produced at least one gesture when completing this test.

#### 3.5.3 Relationship between gesture behavior and evaluation results

Finally, we tested if the gesture behaviors during the learning and then the test phase were related to the results in each evaluation using Spearman correlations. In the Gesture group, no reliable correlation is observed between the number of gestures made during the learning phase and PROP\_CORRECT in Recall, Diagram, Diagram\_Name and QCM, both for short-term and long-term evaluations (r values range from .09 in Recall to .27 in MCQ). This is also the case for QUAL\_EVAL, regardless of the dimension (r values range from -.38 for Engagement to .02 for Clarity). Because about half of the participants didn't make any gestures during the different evaluation tests, it is not possible to reliably correlate the number of gestures made to the scores in the different tests.

## 4 Discussion

The purpose of this study was to provide further evidences that making gestures can improve the acquisition of new knowledge. Anatomy learning was chosen as a specific case of embodied learning and as a challenging topic to learn. In particular, we focused on the specific contribution of making gestures in relation to the learning content as compared with just seeing gestures. Indeed, a stronger

activation of the motor system in the former case could be more beneficial to the reenactment processes supposed to subtend cognitive abilities (embodied cognition, see Barsalou, 2008). To this end, we compared performances in learning anatomy of the forearm between two groups of novice learners: a Gesture group who reproduces the gestures made by a model during learning and a Control group who only see these gestures. Overall, the Gesture group appears to perform better than the Control group but the differences were significant only in the long-term subset of evaluations. Short-term subjective evaluations suggest that the Gesture group felt more engaged in the lecture than the Control group. We also observed a greater tendency to make gestures during MCQ functional evaluation than recall or structural assessments, with yet large between-participants variability. These results are discussed in respect to previous empirical and theoretical work, experimental limits and new perspective.

#### 4.1 Making gestures vs. seeing gestures

Our study investigated the effect of a direct implication of the motor system when learning structural and functional anatomy. The results suggest that “doing gestures” is more effective than just “seeing gestures” while learning the names of anatomical structures. Such effect was previously observed at short-term in a mental rotation task (Goldin-Meadow et al., 2012) and is consistent with the idea that “motor information accrued by the body can affect learning and development by grounding mental representations in motor areas of the cortex and structuring associated perception” (Kontra, Goldin-Meadow, & Beilock, 2012; p732).

In most of the previous work about the role of gestures in learning, “doing gestures” usually overlap “seeing gestures”, meaning the participants who don’t reproduce the gestures don’t see them neither (Cook & Goldin-Meadow, 2006; Cook, Mitchell, & Goldin-Meadow, 2008). If the literature has highlighted that seeing the teacher’s gestures is already helpful for learning (Cook et al., 2013), being active during learning is an important element in the research about learning and memory: Should learners gesture themselves or is seeing the teacher’s gestures enough?

Numerous studies have shown that being active during learning lead to better recognition (Liu, Ward, & Markall, 2007; Meijer & Van der Lubbe, 2011) or faster recognition (Harman, Humphrey, & Goodale, 1999; James, Humphrey, & Goodale, 2001). This effect is observed in different tasks such as matching of objects (Sasaoka, Asakura, & Kawahara, 2010) or textures discrimination (Lee & Wallraven, 2013). The current study provides supplementary evidence in this direction. One could argue that in our study, making gestures may have increase participants’ attention to the lecture and/or the gestures, leading to a better representation of content and gestures. Evidence suggesting that action itself rather than attention supports active learning is provided by Meijer & Van der Lubbe (2011): an active exploration of 3D objects leads to a better recognition of these objects than a passive exploration, even if participants had to carry out a secondary task during learning.

Previous work also suggests that being active during learning of anatomy improve the feeling of involvement, especially when innovative methods are used (cf. Suga et al., 2010). In our study the participants who made gestures tend to evaluate the lecture as more engaging than the participants who did not make gestures. There was no further significant difference for the other qualitative evaluations (clarity, motivation, interactivity, usefulness of gestures). It is possible that the chosen questions were too broad to catch differences between groups or that “seeing gestures” is indeed equivalent as “doing gesture” at a conscious level. It is also possible that novice participants are not able to distinguish the impact of the gestures, as they don’t have reference point of anatomy lecture. Comparisons with students in medical school or physiotherapy will be required to assess this point.

An original aspect of our approach was also to report the number of gestures spontaneously made by the participants during the short-term tests. The quantity of gestures made was limited overall and was clearly participant-specific. Some of the participants might be uncomfortable with making gestures especially when videotaped or observed by someone else, as it was the case in the current study. It will be interesting to compare participants from different cultures and/or ran specific studies to test the interaction between experimental conditions and participants' profiles as gesturers. In addition, our analyses suggest that about half of the participants spontaneously gestured when performing the MCQ test. This test is closer to problem-solving situations, requires functional dynamic knowledge and probably relies more on visuo-spatial and body representations than recall and diagram tests. Further work should analyses long-term use and effect of gestures on performance in functional vs. structural anatomy tests. Investigations of brain activity during testing phase would also allow more direct assessment of the involvement of the motor system in the different evaluations: some participants in the Gesture group may not gesture but their motor system might still be stronger activated than that of participants in the Control group.

## 4.2 Short-term vs. long term performances

An important result of this study is the significant effect of making gestures on performance in long-term assessment only. Furthermore, the participants in the Gesture group were not only better than those of the Control group: they also significantly improved their scores as compared with short-term assessment. Similar beneficial effects of gestures on long-term learning was previously observed in Cook et al., (2008) and Cook, Yip, & Goldin-Meadow (2010) and could be related to effects of sleep on memory. Literature broadly reported positive effect of sleep on memory consolidation (Diekelmann & Born, 2010; Rasch & Born, 2013). During sleep, the neuronal networks involved during learning are reactivated (Ji & Wilson, 2007; McNaughton, 1994) and there is a dialogue between hippocampus and neocortex that consolidates memories (Buzsáki, 1996; Landmann et al., 2014). Therefore, one or two nights of sleep could enhance memory traces by a (re)-simulation of experience (reactivation in the hippocampus) and a reorganization of the memory traces (by the dialogue between hippocampus and neocortex).

The stronger motor activation when making gestures as compared to seeing gestures during learning may influence the reactivation of experience and the reorganization of memory during sleep, favoring the integration of the motor component. Sleep has a beneficial effect both on declarative and procedural memory (Plihal & Born, 1997; Rasch & Born, 2013). Sleep is for example well known to have a positive impact on procedural memory and visuo-spatial tasks (Fischer, Drosopoulos, Tsen, & Born, 2006; Peigneux et al., 2004; Björn Rasch, Büchel, Gais, & Born, 2007), as well as in motor performance (Huber, Felice Ghilardi, Massimini, & Tononi, 2004). The difference between the two groups also suggests that “seeing gestures” vs. “doing gestures” create different memory traces.

We were expecting beneficial effects of making gestures at short-term. We also added the long-term assessment in relation to previous work showing long-term effect of gestures on learning (Cook et al., 2008). The long-term assessment was performed by phone. This limited testing possibilities but reduced the loss of participants, which was important to preserve balance between groups. Both groups went through the same procedure: there was no reason to believe that participants in one group would respect differently the instructions than those in the other group. Regardless of these limits, we can make two important points relative to short-term vs. long-term assessment: (1) the positive effect of making gestures could be present at short-term but hard to observe due to between-subjects design and to the multi-dimensional competences involved in complex learning such as anatomy; (2) longitudinal assessments should be planned systematically in learning studies,



especially when comparing learning conditions, as some conditions may need sleep and/or time to be integrated. This aspect is of particular interest because in the case of learning, long-term retrieval is essential in regards to theoretical aspects as well as practical reasons. Our study is a first step in that way, showing that involving the body can be especially useful to build long-term knowledge.

#### 4.2.1 Why/how doing gestures could improve long-term performances in recall/diagram in an embodied cognition perspective?

In terms of cognitive processes, we can assume that doing gestures during learning is a way to enhance the memory traces associated with learning. According to embodied cognition theories (Barsalou, 2008; Wilson, 2002), cognition emerges in the interaction between the body, the mind and the environment. Memory is defined by every sensory and motor component activated while experiencing the events from daily life (Logan, 1988; Versace et al., 2014). Several memory models are developed according to this idea and usually describe that the number of components of a memory trace can influence its strength (Versace et al., 2014). We can consider that having a supplemental motor component during learning (for the Gesture group) can strengthen the memory trace associated to this learning, if there is enough time (and/or sleeping processing) between learning and retrieval. The embodied cognition theories also suggest that memory is in fact a simulation of previous events (Barsalou, 2009; Barsalou, Pecher, Zeelenberg, Simmons, & Hamann, 2005). Adding a motor component to learning may facilitate the simulation process. Further work should address this hypothesis more directly, in particular using neurophysiological recording during the evaluation phases.

Anatomy of the arm was chosen as topic as a specific case of embodied learning: gestures involved in learning were directly related to the learning content and participants could represent the concepts and vocabulary in relation to their own body. Moreover, the upper limb is visible and this information can be used as clue for retrieval, which is not the case for anatomy of the face for example (Dickson & Stephens, 2015). An opened question is then: Is the beneficial effect of gestures related to the type of link between the gestures and knowledge to learn? Would the benefit from gestures be equivalent for another topic, unrelated to the body? Would less specific gestures (e.g. abstract gestures) lead to the same effects? Gestures seem to be useful in different domains (Susan Goldin-Meadow & Alibali, 2013; McCafferty, 2004; Novack & Goldin-Meadow, 2015; Trofatter et al., 2014) but the relevance of the link between learning content and gestures type remains an opened question. In our study, participants who gestures during testing mainly do so during MCQ functional test. Understanding the embodiment of learning through gesturing will require investigating more specifically the link between the gestures types and the nature of learning itself. Anatomy learning could be used to manipulate this link and provide a major contribution for embodiment theories and educational approaches. Our study was a first step in that direction.

## 5 References

- Abner, N., Cooperrider, K., & Goldin-Meadow, S. (2015). Gesture for Linguists: A Handy Primer. *Linguistics and Language Compass*, 9(11), 437–451. <http://doi.org/10.1111/lnc3.12168>
- Alibali, M. W., & Nathan, M. J. (2012). Embodiment in Mathematics Teaching and Learning: Evidence From Learners' and Teachers' Gestures. *Journal of the Learning Sciences*, 21(2), 247–286. <http://doi.org/10.1080/10508406.2011.611446>
- Barsalou, L. (2008). Grounded Cognition. *Annual Review of Psychology*, 59(1), 617–645. <http://doi.org/10.1146/annurev.psych.59.103006.093639>

- 631 Barsalou, L. (2009). Simulation, situated conceptualization, and prediction. *Philosophical*  
632 *Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1281–1289.  
633 <http://doi.org/10.1098/rstb.2008.0319>
- 634 Barsalou, L. W., Pecher, D., Zeelenberg, R., Simmons, W. K., & Hamann, S. B. (2005). Multimodal  
635 Simulation in Conceptual Processing. *Categorization inside and Outside the Lab: Essays in*  
636 *Honor of Douglas L. Medin*. Retrieved from  
637 [http://psychology.emory.edu/cognition/barsalou/papers/Barsalou\\_et\\_al\\_chap\\_2005\\_multimodal](http://psychology.emory.edu/cognition/barsalou/papers/Barsalou_et_al_chap_2005_multimodal_simulation.pdf%5Cnpapers3://publication/uuid/1DFF0034-1441-48CB-8E71-AB349B87B0DB)  
638 [\\_simulation.pdf%5Cnpapers3://publication/uuid/1DFF0034-1441-48CB-8E71-](http://psychology.emory.edu/cognition/barsalou/papers/Barsalou_et_al_chap_2005_multimodal_simulation.pdf%5Cnpapers3://publication/uuid/1DFF0034-1441-48CB-8E71-AB349B87B0DB)  
639 [AB349B87B0DB](http://psychology.emory.edu/cognition/barsalou/papers/Barsalou_et_al_chap_2005_multimodal_simulation.pdf%5Cnpapers3://publication/uuid/1DFF0034-1441-48CB-8E71-AB349B87B0DB)
- 640 Beaudoin-Ryan, L., & Goldin-Meadow, S. (2014). Teaching moral reasoning through gesture.  
641 *Developmental Science*, 17(6), 984–990. <http://doi.org/10.1111/desc.12180>
- 642 Berney, S., Bétrancourt, M., Molinari, G., & Hoyek, N. (2015). How spatial abilities and dynamic  
643 visualizations interplay when learning functional anatomy with 3D anatomical models.  
644 *Anatomical Sciences Education*, 8(5), 452–462. <http://doi.org/10.1002/ase.1524>
- 645 Borghi, A. M., & Pecher, D. (2011). Introduction to the Special Topic Embodied and Grounded  
646 Cognition. *Frontiers in Psychology*, 2(August), 2010–2012.  
647 <http://doi.org/10.3389/fpsyg.2011.00187>
- 648 Buzsáki, G. (1996). The hippocampo-neocortical dialogue. *Cereb Cortex*, 6(2), 81–92.  
649 <http://doi.org/10.1093/cercor/6.2.81>
- 650 Casasanto, D., & Dijkstra, K. (2010). Motor action and emotional memory. *Cognition*, 115(1), 179–  
651 185. <http://doi.org/10.1016/j.cognition.2009.11.002>
- 652 Christensen, R. H. B. (2015). ordinal---Regression Models for Ordinal Data. Retrieved from  
653 <http://www.cran.r-project.org/package=ordinal/>
- 654 Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The Role of Gesture in Bilingual  
655 Education: Does Gesture Enhance Learning? *International Journal of Bilingual Education and*  
656 *Bilingualism*, 7(4), 303–319. <http://doi.org/10.1080/13670050408667815>
- 657 Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3(9), 345–351.  
658 [http://doi.org/10.1016/S1364-6613\(99\)01361-3](http://doi.org/10.1016/S1364-6613(99)01361-3)
- 659 Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after  
660 observing hand gesture. *Child Development*, 84(6), 1863–1871.  
661 <http://doi.org/10.1111/cdev.12097>
- 662 Cook, S. W., & Goldin-Meadow, S. (2006). The Role of Gesture in Learning: Do Children Use Their  
663 Hands to Change Their Minds? *Journal of Cognition and Development*, 7(2), 211–232.  
664 [http://doi.org/10.1207/s15327647jcd0702\\_4](http://doi.org/10.1207/s15327647jcd0702_4)
- 665 Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*,  
666 106(2), 1047–58. <http://doi.org/10.1016/j.cognition.2007.04.010>
- 667 Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal*

- 668 *of Memory and Language*, 63(4), 465–475. <http://doi.org/10.1016/j.jml.2010.07.002>
- 669 Dickson, K. A., & Stephens, B. W. (2015). It's all in the mime: Actions speak louder than words  
670 when teaching the cranial nerves. *Anatomical Sciences Education*, 8, 584–592.  
671 <http://doi.org/10.1002/ase.1531>
- 672 Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*,  
673 11(2), 114–126. <http://doi.org/10.1038/nrn2762>
- 674 Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit learning -- explicit knowing: a role  
675 for sleep in memory system interaction. *Journal of Cognitive Neuroscience*, 18(3), 311–319.  
676 <http://doi.org/10.1162/jocn.2006.18.3.311>
- 677 Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin &*  
678 *Review*, 9(3), 558–65. <http://doi.org/http://dx.doi.org/10.3758/bf03196313>
- 679 Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating  
680 language. *Annual Review of Psychology*, 64, 257–283. <http://doi.org/10.1146/annurev-psych-113011-143802>
- 682 Goldin-Meadow, S., & Beilock, S. L. (2010). Action's Influence on Thought: The Case of Gesture.  
683 *Perspectives on Psychological Science*, 5(6), 664–674.  
684 <http://doi.org/10.1177/1745691610388764>
- 685 Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind  
686 about math. *Journal of Educational Psychology*, 91(4), 720–730. <http://doi.org/10.1037/0022-0663.91.4.720>
- 688 Goldin-Meadow, S., Levine, S. C., Zinchenko, E., Yip, T. K., Hemani, N., & Factor, L. (2012).  
689 Doing gesture promotes learning a mental transformation task better than seeing gesture.  
690 *Developmental Science*, 15(6), 876–884. <http://doi.org/10.1111/j.1467-7687.2012.01185.x>
- 691 Gullberg, M. (2006). Some reasons for studying gesture and second language acquisition (Homage  
692 à Adam Kendon). *IRAL - International Review of Applied Linguistics in Language Teaching*,  
693 44(2), 103–124. <http://doi.org/10.1515/IRAL.2006.004>
- 694 Gunderson, E. A., Spaepen, E., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2015). Gesture as a  
695 window onto children's number knowledge. *Cognition*, 144, 14–28.  
696 <http://doi.org/10.1016/j.cognition.2015.07.008>
- 697 Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views  
698 facilitates visual recognition. *Current Biology*, 9(22), 1315–1318.  
699 [http://doi.org/http://dx.doi.org/10.1016/S0960-9822\(00\)80053-6](http://doi.org/http://dx.doi.org/10.1016/S0960-9822(00)80053-6)
- 700 Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action.  
701 *Psychonomic Bulletin & Review*, 15(3), 495–514. <http://doi.org/10.3758/PBR.15.3.495>
- 702 Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models.  
703 *Biometrical Journal*, 50(3), 346–363.

- 704 Hoyek, N., Collet, C., Rastello, O., Fargier, P., Thiriet, P., & Guillot, A. (2009). Enhancement of  
705 mental rotation abilities and its effect on anatomy learning. *Teaching and Learning in Medicine*,  
706 21(3), 201–6. <http://doi.org/10.1080/10401330903014178>
- 707 Hu, F. T., Ginns, P., & Bobis, J. (2015). Getting the point: Tracing worked examples enhances  
708 learning. *Learning and Instruction*, 35, 85–93. <http://doi.org/10.1016/j.learninstruc.2014.10.002>
- 709 Huber, R., Felice Ghilardi, M., Massimini, M., & Tononi, G. (2004). Local sleep and learning.  
710 *Nature*, 430(6995), 78–81. <http://doi.org/10.1038/nature02663>
- 711 Ionescu, T., & Vasc, D. (2014). Embodied Cognition: Challenges for Psychology and Education.  
712 *Procedia - Social and Behavioral Sciences*, 128, 275–280.  
713 <http://doi.org/10.1016/j.sbspro.2014.03.156>
- 714 Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development.  
715 *Psychological Science*, 16(5), 367–371. <http://doi.org/10.1111/j.0956-7976.2005.01542.x>
- 716 James, K. H., Humphrey, G. K., & Goodale, M. a. (2001). Manipulating and recognizing virtual  
717 objects: where the action is. *Canadian Journal of Experimental Psychology = Revue*  
718 *Canadienne de Psychologie Experimentale*, 55(1961), 111–120.  
719 <http://doi.org/10.1037/h0087358>
- 720 Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus  
721 during sleep. *Nature Neuroscience*, 10(1), 100–107. <http://doi.org/10.1038/nn1825>
- 722 Kelly, S. D., Manning, S. M., & Rodak, S. (2008). Gesture gives a hand to language and learning:  
723 Perspectives from cognitive neuroscience, developmental psychology and education. *Linguistics*  
724 *and Language Compass*, 2(4), 569–588. <http://doi.org/10.1111/j.1749-818X.2008.00067.x>
- 725 Kontra, C., Goldin-Meadow, S., & Beilock, S. L. (2012). Embodied Learning Across the Life Span.  
726 *Topics in Cognitive Science*, 4(4), 731–739. <http://doi.org/10.1111/j.1756-8765.2012.01221.x>
- 727 Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalter, K., ... Nissen, C.  
728 (2014). The reorganisation of memory during sleep. *Sleep Medicine Reviews*, 18(6), 531–541.  
729 <http://doi.org/10.1016/j.smrv.2014.03.005>
- 730 Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES--ELAN system.  
731 *Behavior Research Methods*, 41(3), 841–9. <http://doi.org/10.3758/BRM.41.3.841>
- 732 Lee, H., & Wallraven, C. (2013). Exploiting object constancy: Effects of active exploration and  
733 shape morphing on similarity judgments of novel objects. *Experimental Brain Research*, 225(2),  
734 277–289. <http://doi.org/10.1007/s00221-012-3368-1>
- 735 Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals.  
736 *Soviet Physics Doklady*. <http://doi.org/citeulike-article-id:311174>
- 737 Liu, C. H., Ward, J., & Markall, H. (2007). The role of active exploration of 3D face stimuli on  
738 recognition memory of facial information. *Journal of Experimental Psychology. Human*  
739 *Perception and Performance*, 33(4), 895–904. <http://doi.org/10.1037/0096-1523.33.4.895>

- 740 Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4),  
741 492–527. <http://doi.org/10.1037/0033-295X.95.4.492>
- 742 Lorant, J., & Nicolas, A. (2004). Validation de la traduction française du Movement Imagery  
743 Questionnaire-Revised (MIQ-R); Validation of the French translation of the Movement Imagery  
744 Questionnaire-Revised (MIQ-R). *Science et Motricité Paris*, (53), 57–68.  
745 <http://doi.org/10.3917/sm.053.0057>
- 746 Macken, L., & Ginns, P. (2014). Pointing and tracing gestures may enhance anatomy and physiology  
747 learning. *Medical Teacher*, 36(7), 596–601. <http://doi.org/10.3109/0142159X.2014.899684>
- 748 McCafferty, S. G. (2004). Space for cognition: gesture and second language learning. *International*  
749 *Journal of Applied Linguistics*, 14(1), 148–165. [http://doi.org/10.1111/j.1473-](http://doi.org/10.1111/j.1473-4192.2004.0057m.x)  
750 [4192.2004.0057m.x](http://doi.org/10.1111/j.1473-4192.2004.0057m.x)
- 751 McMenamin, P. G. (2008). Body painting as a tool in clinical anatomy teaching. *Anatomical*  
752 *Sciences Education*, 1(4), 139–144. <http://doi.org/10.1002/ase.32>
- 753 McNaughton, M. A. W. and B. L. (1994). Reactivation of Hippocampal Ensemble Memories During  
754 Sleep. *Science*, 265(July), 5–8. <http://doi.org/10.1126/science.8036517>
- 755 McNeil, N., Alibali, M., & Evans, J. (2000). The role of gesture in children’s comprehension of  
756 spoken language: Now they need it, now they don’t. *Journal of Nonverbal Behavior*, 24(2),  
757 131–150. <http://doi.org/10.1023/A:1006657929803>
- 758 McNeill, D. (1981). Action, thought and language. *Cognition*, 10(1–3), 201–208.  
759 [http://doi.org/10.1016/0010-0277\(81\)90047-0](http://doi.org/10.1016/0010-0277(81)90047-0)
- 760 McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago  
761 press.
- 762 Meijer, F., & Van der Lubbe, R. H. J. (2011). Active exploration improves perceptual sensitivity for  
763 virtual 3D objects in visual recognition tasks. *Vision Research*, 51(23–24), 2431–9.  
764 <http://doi.org/10.1016/j.visres.2011.09.013>
- 765 New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français  
766 contemporain sur internet□: LEXIQUE™//A lexical database for contemporary french□:  
767 LEXIQUE™. *L’année Psychologique*, 101(3), 447–462. <http://doi.org/10.3406/psy.2001.1341>
- 768 Novack, M., & Goldin-Meadow, S. (2015). Learning from Gesture: How Our Hands Change Our  
769 Minds. *Educational Psychology Review*, 27, 405–412. [http://doi.org/10.1007/s10648-015-9325-](http://doi.org/10.1007/s10648-015-9325-3)  
770 [3](http://doi.org/10.1007/s10648-015-9325-3)
- 771 Oh, C. S., Won, H. S., Kim, K. J., & Jang, D. S. (2011). “Digit anatomy”: A new technique for  
772 learning anatomy using motor memory. *Anatomical Sciences Education*, 4(3), 132–141.  
773 <http://doi.org/10.1002/ase.223>
- 774 Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., ... Maquet, P. (2004). Are  
775 spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron*,  
776 44(3), 535–545. <http://doi.org/10.1016/j.neuron.2004.10.007>

- 777 Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn  
778 Vandenberg and Kuse mental rotations test: different versions and factors that affect  
779 performance. *Brain and Cognition*, 28(1), 39–58. <http://doi.org/10.1006/brcg.1995.1032>
- 780 Ping, R. M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to  
781 teach children conservation of quantity. *Developmental Psychology*, 44(5), 1277–1287.  
782 <http://doi.org/10.1037/0012-1649.44.5.1277>
- 783 Plihal, W., & Born, J. (1997). Effects of Early and Late Nocturnal Sleep on Declarative and  
784 Procedural Memory. *Journal of Cognitive Neuroscience*, 9(4), 534–547.  
785 <http://doi.org/10.1162/jocn.1997.9.4.534>
- 786 R Core Team. (2015). *R: A language and Environment for Statistical Computing*.
- 787 R Development Core Team. (2016). *R: A Language and Environment for Statistical Computing*. *R*  
788 *Foundation for Statistical Computing Vienna Austria*. <http://doi.org/10.1038/sj.hdy.6800737>
- 789 Rasch, B., & Born, J. (2013). About Sleep’s Role in Memory. *Physiological Reviews*, 93(2), 681–  
790 766. <http://doi.org/10.1152/physrev.00032.2012>
- 791 Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt  
792 declarative memory consolidation. *Science (New York, N.Y.)*, 315(5817), 1426–9.  
793 <http://doi.org/10.1126/science.1138581>
- 794 Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of  
795 lexical movements in speech production. *Psychological Science*, 7(4), 226.  
796 <http://doi.org/10.1111/j.1467-9280.1996.tb00364.x>
- 797 Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in Early Gesture Explain SES Disparities in  
798 Child Vocabulary Size at School Entry. *Science*, 323(5916), 951–953.  
799 <http://doi.org/10.1126/science.1167025>
- 800 Sasaoka, T., Asakura, N., & Kawahara, T. (2010). Effect of active exploration of 3-D object views on  
801 the view-matching process in object recognition. *Perception*, 39(3), 289–308.  
802 <http://doi.org/10.1068/p5721>
- 803 Skinder-Meredith, A. E. (2010). Innovative activities for teaching anatomy of speech production.  
804 *Anatomical Sciences Education*, 3(5), 234–43. <http://doi.org/10.1002/ase.173>
- 805 Skulmowski, A., Pradel, S., Kühnert, T., Brunnett, G., & Rey, G. D. (2016). Embodied learning  
806 using a tangible user interface: The effects of haptic perception and selective pointing on a  
807 spatial learning task. *Computers and Education*, 92–93, 64–75.  
808 <http://doi.org/10.1016/j.compedu.2015.10.011>
- 809 Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression  
810 with beta-distributed dependent variables. *Psychological Methods*, 11(1), 54–71.  
811 <http://doi.org/10.1037/1082-989X.11.1.54>
- 812 Sugand, K., Abrahams, P., & Khurana, A. (2010). The anatomy of anatomy: A review for its  
813 modernization. *Anatomical Sciences Education*, (MARCH 2010), NA-NA.



- 814 <http://doi.org/10.1002/ase.139>
- 815 Trofatter, C., Kontra, C., Beilock, S., & Goldin-Meadow, S. (2014). Gesturing has a larger impact on  
816 problem-solving than action, even when action is accompanied by words. *Language, Cognition*  
817 *and Neuroscience, In Press*(0), 1–10. <http://doi.org/10.1080/23273798.2014.905692>
- 818 Tyler, L. K., Stamatakis, E. a., Dick, E., Bright, P., Fletcher, P., & Moss, H. (2003). Objects and their  
819 actions: Evidence for a neurally distributed semantic system. *NeuroImage*, 18(2), 542–557.  
820 [http://doi.org/10.1016/S1053-8119\(02\)00047-2](http://doi.org/10.1016/S1053-8119(02)00047-2)
- 821 Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning:  
822 A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), 187–204.  
823 [http://doi.org/10.1016/S0361-476X\(02\)00007-3](http://doi.org/10.1016/S0361-476X(02)00007-3)
- 824 Varela, F. J., Thompson, E., & Rosch, E. (1993). *L'inscription corporelle de l'esprit : Sciences*  
825 *cognitives et expérience humaine*. Seuil.
- 826 Versace, R., Vallet, G. T., Riou, B., Lesourd, M., Labeye, E., & Brunel, L. (2014). Act-In: An  
827 integrated view of memory mechanisms. *Journal of Cognitive Psychology*, 26(3), 280–306.  
828 <http://doi.org/10.1080/20445911.2014.892113>
- 829 Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*,  
830 68(1), 77–94. [http://doi.org/10.1016/S0010-0277\(98\)00032-8](http://doi.org/10.1016/S0010-0277(98)00032-8)
- 831 Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–  
832 636. <http://doi.org/10.3758/BF03196322>
- 833 Zumwalt, A. C., Lufler, R. S., Monteiro, J., Shaffer, K., Luefler, R. S., Monteiro, J., & Shaffer, K.  
834 (2010). Building the body: Active learning laboratories that emphasize practical aspects of  
835 anatomy and integration with radiology. *Anatomical Sciences Education*, 3(3), 134–140.  
836 <http://doi.org/10.1002/ase.153>

## 837 **6 Conflict of Interest**

838 The authors declare that the research was conducted in the absence of any commercial or financial  
839 relationships that could be construed as a potential conflict of interest.

## 840 **7 Author Contributions**

841 MC, ARC, JT and OP: designed the experiment, interpreted the results

842 MC and OP: created the material for the experiment

843 MC: performed the experiment

844 MC, ARC and SG: analyzed the data

845 MC and ARC: wrote the manuscript

846 JT, OP and SG: performed revisions and approved for the manuscript for publication

## 847 **8 Funding**

848 This work has been supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) and by the  
849 Pôle Grenoble Cognition.

## 850 **9 Acknowledgments**

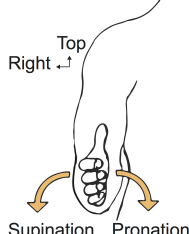
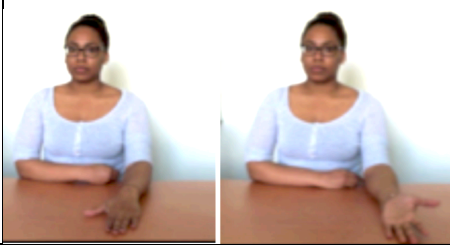
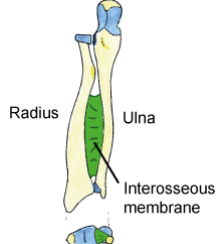
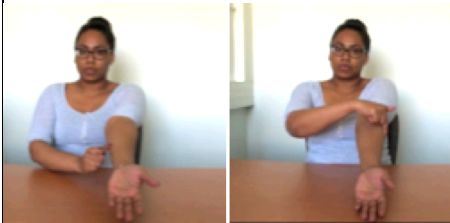
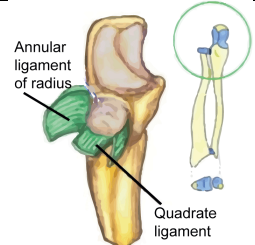

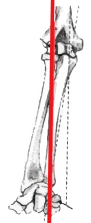
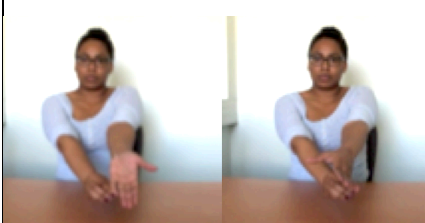
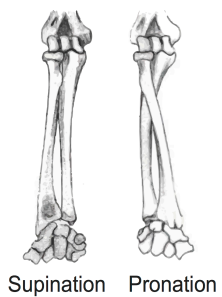
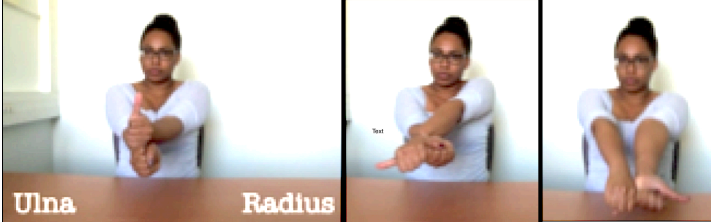
851 The authors are grateful to their colleagues from Gipsa-lab for their support and to the study  
852 participants.

853

854 **10 Tables of the manuscript**

855 Table 1: Main steps of the lesson, associated diagrams and gestures made by the model. The left  
 856 column illustrates the principal diagrams displayed in the different slides. The right column provides  
 857 a short description of the educational content and shots from the video illustrating the types of  
 858 gestures made by the model. Note that the model mirrors the diagrams. The number of gestures and  
 859 slides are also specified for each step.

860

Diagram	Gestures and content	
 <p>Top Right</p> <p>Supination Pronation</p>		<p>Step 1 (2 gestures, 1 slide). Supination and pronation movements of the forearm are introduced. The model makes pronation (left) and supination (right) rotating motions from the reference position (diagram on the left). Video shots display movements offset positions.</p>
 <p>Radius Ulna</p> <p>Interosseous membrane</p>		<p>Step 2 (3 gestures, 4 slides). The structures of the forearm are explained (the original diagram includes more details presented in several steps, see Figure 1). Here, the model shows on her arm the location of the structures by moving her finger along her ulna (left) and her radius (right) bones. She then did the same for the interosseous membrane</p>
 <p>Annular ligament of radius</p> <p>Quadrate ligament</p>		<p>Step 3 (2 gestures, 1 slide). The main ligaments involved in supination and pronation and their function are explained. The model mimics the rotation of the radial head into the annular ligament of the ulna: she brings her right fist into her left hand and then execute rotational movements of her fist. A similar gesture was used for the second ligament : the quadrate ligament.</p>
		<p>Step 4 (1 gesture, 1 slide). The rotation axis of pronation and supination is explained. The model shows the axis from the radial head to the top of her pinky finger. She then pivots her arm around this axis.</p>
 <p>Supination Pronation</p>	 <p>Ulna Radius</p>	<p>Step 5 (2 gestures, 1 slide). The function of the radius and ulna during pronation and supination are explained. The model associates one of her arms to the radius (thumb-up arm) and the other to the ulna. Form left to right: in the neutral position (see diagram on row 1), the radius is on the top, it then pivots over the ulna in pronation while staying almost parallel to it in supination.</p>

861

862

Table 2: Group description in terms of visuo-spatial abilities (MRT, MIQ), memory (Words), age, educational level and number of days between learning and long-term assessment.

	<b>Control (n=21)</b>	<b>Gesture (n=21)</b>
	<b>Mean <math>\pm</math> SE</b>	<b>Mean <math>\pm</math> SE</b>
<b>MRT</b>	9.1 $\pm$ 0.69	8.9 $\pm$ 0.82
<b>MIQ</b>	42.3 $\pm$ 1.26	43.9 $\pm$ 1.23
<b>Words</b>	4.8 $\pm$ 0.36	4.8 $\pm$ 0.34
<b>Age (years)</b>	25.6 $\pm$ 1.23	27.7 $\pm$ 1.50
<b>Educational Level</b> (Nb. of years of studies after high school diploma)	3.1 $\pm$ 0.48	3.5 $\pm$ 0.51
<b>Nb. of days</b>	2.4 $\pm$ 0.18	2.6 $\pm$ 0.15

## 11 Figure legends

Figure 1. Main diagram of the forearm anatomy used in learning and testing phases. In total 14 different names had to be memorized on the main diagram (left) and two more on the diagram with ligaments (on the right). The two diagrams share two structures with different views (1 and 15; 2 and 18).

Figure 2. Overview of the procedure of the experiment (see text for details)

Figure 3. Average proportion of correct responses (PROP\_CORRECT) in each test and group in the short-term session (n=21 in each group). Error-bars represent between participants standard errors.

Figure 4. Average proportion of correct responses (PROP\_CORRECT) for the participants and the tests performed in both the short-term and long-term session (n=19 in each group). Error-bars represent between participants standard errors.

Figure 5. Distribution of QUAL\_EVAL scores (subjective evaluations on Likert-scale, 10: most positive evaluation) for each DIMENSION depending on the GROUP.

Figure 6. Number of gestures made by each participant during the learning phase (first row) and the evaluation phase (second row). Results are given for each step of the learning phase and each test in the evaluation phase. The number of expected gestures was: 2 gestures in Steps 1, 3 and 5; 3 gestures in Step 2; 1 gesture in Step 4, see Table 1 for details.