The revised NEUROGES–ELAN system: An objective and reliable interdisciplinary analysis tool for nonverbal behavior and gesture

Hedda Lausberg[1] · Han Sloetjes[2]

Abstract As visual media spread to all domains of public and scientific life, nonverbal behavior is taking its place as an important form of communication alongside the written and spoken word. An objective and reliable method of analysis for hand movement behavior and gesture is therefore currently required in various scientific disciplines, including psychology, medicine, linguistics, anthropology, sociology, and computer science. However, no adequate common methodological standards have been developed thus far. Many behavioral gesture-coding systems lack objectivity and reliability, and automated methods that register specific movement parameters often fail to show validity with regard to psychological and social functions. To address these deficits, we have combined two methods, an elaborated behavioral coding system and an annotation tool for video and audio data. The NEUROGES–ELAN system is an effective and user-friendly research tool for the analysis of hand movement behavior, including gesture, self-touch, shifts, and actions. Since its first publication in 2009 in *Behavior Research Methods*, the tool has been used in interdisciplinary research projects to analyze a total of 467 individuals from different cultures, including subjects with mental disease and brain damage. Partly on the basis of new insights from these studies, the system has been revised methodologically and conceptually. The article presents the revised version of the system, including a detailed study of reliability. The improved reproducibility of the revised version makes NEUROGES–ELAN a suitable system for basic empirical research into the relation between hand movement behavior and gesture and cognitive, emotional, and interactive processes and for the development of automated movement behavior recognition methods.

The NEUROGES–ELAN System is an analysis tool for nonverbal behavior focusing on body movement and gesture. Apart from behavior analysis per se, it is suited for basic research on cognitive, emotional, and interactive processes via analyzing nonverbal behavior. Since it enables the investigation of processes that are not verbalized, its specific potential lies in the exploration of implicit cognitive, emotional, and interactive processes that may be conducted beyond awareness. The fields of application include psychology, neuropsychology, medicine, evolutionary anthropology, linguistics, and related areas. Furthermore, because NEUROGES-ELAN offers the option to analyze gestures independently of speech, it is suited for linguistic research testing basic paradigms in the relation between speech and gesture. As a highly operationalized analysis tool, it has also been used for developing automatic movement recognition methods. Clinical fields of application are doctor–patient interaction, psychodiagnostics, and consultation/therapy process and outcome control.

NEUROGES–ELAN combines the behavioral analysis system NEUROGES with the annotation tool for video and audio data ELAN. The extended version of the NEUROGES system (Lausberg, 2013) is designed for the analysis of nonverbal behavior including head, trunk, hand/arm/shoulder, and

✉ Hedda Lausberg
neuropsychosomatik@dshs-koeln.de

1 Department of Neurology, Psychosomatic Medicine, and Psychiatry, German Sport University Cologne, Am Sportpark Müngersdorf 6, D-50939 Köln, Germany

2 The Language Archive, Max Planck Institute for Psycholinguistics, Wundtlaan 1, Nijmegen 6525 XD, The Netherlands

🙋 Springer

foot/leg movements. However, among these four subsystems the hand/arm/shoulder subsystem, which has been published in BRM in 2009, has been used most widely. In seven steps comprising coding algorithm, the ongoing stream of hand/arm/shoulder movements (hereafter, "hand movements") is segmented and classified into more and more fine-grained movement units. At each assessment step (representing a category), specific movement criteria are applied in order to segment the behavior and to classify the resulting units with values. Notably, the choice of movement criteria is based on neuropsychological and psychological research. The seven assessment steps are grouped in three modules: Module I (Steps 1–3) deals with aspects of hand movement behavior that are related to specific neuropsychological processes. For example, the Structure category (Step 2) provides information about conceptualization processes by analyzing the trajectory of the movement, and the Focus category (Step 3) refers to attention processes by analyzing the location where the hand acts. Module II (Steps 4–5) focuses on the laterality of hand movement behavior, including complex aspects such as dominance. It thereby addresses questions of hemispheric specialization and inter-hemispheric cooperation. Module III (Steps 6–7) analyzes the function of hand movements, which includes the analysis of the meaning of gestures. Notably,—just as in the Modules I and II—the Module III analysis is based on the visual appearance of the movement only, that is, it refers to those aspects of the function of a hand movement that are pre-determined by its temporo-spatial form. At each assessment step, if it applies, subunits can be generated. Since the revised version of ELAN enables the concatenation of all previous assessments (values of the steps 1 - 5) for each hand movement unit, Module III analysis starts with fine-grained hand movement units that are precisely pre-classified according to trajectory, location, laterality, and so forth. These units are specified by further movement criteria that are, among others, suited to determine the meaning of gestures based on their temporo–spatial form.

The coding algorithm and the precise definitions of the movement criteria and the values are described in 200 pages comprising coding manual (available from the first author). For its application with ELAN (https://tla.mpi.nl/tools/tla-tools/elan/) the NEUROGES coding sheet has been transformed into an ELAN template file (www.neuroges-bast.info). The video that shall be analysed is linked with the NEUROGES–ELAN template and then the behavior is segmented by tagging units and annotating them with a value.

An extensive review of the existing systems for the analysis of hand movement behavior and gesture preceded the development of the NEUROGES system. Descriptive movement values that prove to be valid with regard to cognitive, emotional, and interactive processes have been adopted in NEUROGES. Common methodological shortcomings of the existing systems have been remedied. Furthermore, empirical findings from expression psychology on how observers interpret movement behavior have been considered. Approaches from psychotherapy research on pattern detection in movement behavior have further influenced the NEUROGES design. Finally, the system has been created on the basis of recent neuropsychological findings that underline the relevance of hand movement laterality for exploring the relationship between hand movement behavior including gesture and cognitive, emotional, and interactive functions.

The bidirectional link between movement form and movement function

A variety of coding systems are currently available for hand movement behavior and gesture research—for example, Efron (1941), Ekman and Friesen (1969), Freedman (1972), Kimura (1973a, b), or McNeill (1992). Many of these systems operate with hand movement or gesture values that are defined by their function—for example, "Regulators… These are acts which maintain and regulate the back-and-forth nature of speaking and listening between two or more interactants" (Ekman & Friesen, 1969, p. 82). Precise descriptions of the visual appearance of the movement—for example, how the hand moves or how it is shaped—are rarely provided.

Many of these function-oriented hand movement and gesture coding systems ignore the bidirectional link between movement function and movement form. For instance, in Ekman and Friesen's (1969) coding system, each of the functionally defined movement values may be represented by a broad range of diverse movement forms. As an example, *regulators* can be position shifts but also head nods. The implicit paradigm behind this methodological design is that any function can be associated with any form, and vice versa.

However, neuropsychological research strongly challenges this paradigm, as it evidences that the production of specific movement forms are associated with specific cognitive functions. As an example, the production of *hand-head positions*, in which the hand adopts a specific orientation relative to the head such as in the military salute, is lateralized to the left hemisphere, whereas the generation of complex *finger configurations*, in which the fingers of a hand form a complex spatial configuration such in the Peace sign or the O.K. sign, relies on additional right-hemispheric competences (Goldenberg, 1996, 1999; Goldenberg, Laimgruber, & Hermsdörfer, 2001; Kimura & Archibald, 1974; Lausberg & Cruz, 2004). Goldenberg proposes that the generation of *hand–head positions* puts greater demands on body part coding, which is a left hemispheric function, whereas the production of complex *finger configuration* requires spatial competences, which are lateralized to the right hemisphere (Goldenberg, 1999; Goldenberg & Strauss, 2002). In another example, *pantomime* gestures in which the hand is shaped as if

it held an imaginary tool—for example, the hand is shaped as if it held a toothbrush—can only be generated in the left hemisphere. This is compatible with the fact that abstract thinking is a left hemispheric competence (e.g., Lezak, 1995). In contrast, *pantomime* gestures in which the hand is used as if it were the tool—for example, the index represents a toothbrush (body-part-as-object BPO)—can also be generated in the right hemisphere (see the review in Lausberg, Cruz, Kita, Zaidel, & Ptito, 2003). Thus, neuropsychological research provides ample evidence that specific movement forms are lateralized to the right and left hemispheres and accordingly, that they are associated with specific cognitive functions.

Furthermore, the function of a movement determines its form. For instance, in order to indicate a precise location in space it is most effective to shape the hand in such a way that it can serve as the starting point of a vector—for example, the extended index finger. The imaginary prolongation of this vector leads to the intended target in space. Floppy hand movements with a relaxed hand, for example, would not be effective for pointing out precise locations in space. Regarding the form to function direction, the form of a movement allows only for a limited set of functions or meanings. For instance, if both hands create the shape of a triangle, this gesture could refer to all concrete and abstract entities that share aspects of a triangle. However, it would not, for instance, serve to refer to round entities. Recent empirical studies using event related potentials indicate that the brain reliably detects incongruencies between the gestural form and the meaning, as conveyed by the word that accompanies the gesture (Kelly, Kravitz, & Hopkins, 2004).

Further empirical evidence that questions function-defined hand movement and gesture coding systems stems from studies that follow the tradition of expression psychology. These studies demonstrate that interpretations of the function and meaning of body movement by untrained observers are often wrong and they cannot serve as a solid basis for empirical research. Wallbott (1989) investigated how untrained observers assessed psychiatric patients' nonverbal behavior. The 20 raters were shown videos of clinical admission and discharge interviews without sound, and they were asked to estimate on the basis of the nonverbal behavior whether the interviews were from admission or from discharge. The raters' admission/discharge attributions turned out to be completely incorrect. Wallbott further examined on which movement criteria (forms) the untrained raters grounded their decisions. In fact, the raters "were not 'wildly guessing'" (Wallbott, 1989, p. 142), but they systematically employed specific movement criteria, such as intensity or expansiveness. Their premises about the relation between movement behavior and admission/discharge, however, were wrong. Wallbott's findings concurred with earlier interpretation experiments that unanimously revealed that untrained raters made incorrect judgments about participants' personalities on the basis of

observation of gait or on photographs (Eisenberg & Reichline, 1939; Mason, 1957). In this line of research, Frijda (1965) underlined that it was not only important to understand the principles of the meaning of expression but also to explore the principles of the assessment of expression.

Moreover, common popular assumptions about the relationship between psychological processes and movement behavior often do not survive under scientific scrutiny. Contrary to their own hypotheses, Allport and Vernon (1933) found no correlation between the tempo of motor actions and the tempo of cognitive operations. Rimoldi (1951), who analyzed 59 motoric and cognitive velocity tests, found no correlation between the parameters of these two groups. The exception is the finding by Eisenberg (1937) who compared speed, expansion in space, and pressure when walking, writing, and drawing in individuals with extremely high and low feelings of self-dominance, as measured with the Social Personality Inventory. Dominant individuals achieved significantly higher scores in all motor parameters than the nondominant ones. Hargadine (1973) found practically no correlation between self-actualization, as registered with the Personal Orientation Inventory (Shostrom, 1972), and the size of the movement repertoire according to Movement Scope Check List. The only significant correlation was the size of movement repertoire with a positive outlook on nature and mankind. Burn (1987) examined in females with anorexia the internal–external locus of control, as measured with the multidimensional scales by Reid and Ware (1974), and the movement behavior, as measured with the Laban movement analysis. The unexpected significant finding was that the subjects had lesser use of near reach space when correlated with an external orientation. Lausberg, von Wietersheim, and Feiereis (1996) examined the relationship between personality, as measured with the Freiburger Personality Inventory (FPI; Fahrenberg, Hampel, & Selg, 1984), and movement behavior, as measured with BAST (Lausberg, 1998), in 120 females. Common assumptions about the relationship between personality and movement behavior could not be affirmed. For instance, the movement parameter *balance* did not correlate with the personality parameter *emotional instability*, and *use of strength* when stamping did not correlate with the degree of *spontaneous aggression*. Instead, the unexpected finding was that three FPI-parameters *nervousness*, *openness*, and *masculinity* correlated with the movement parameter *half of the body* although they did not correlate with each other. A low degree of *nervousness*, a high degree of *openness*, and a high degree of *masculinity* correlated with a preference for moving the *lower half of the body* as compared to the upper half. Thus, studies of the interpretation of movement function by untrained raters and studies on the relationship between movement form and psychological function unequivocally evidence the need for basic research on the relation between movement form and function.

A prerequisite for this kind of basic research is that movement forms are objectively defined and reliably identified behavioral entities. Since hand movements and gestures are primarily perceived as spatio-temporal phenomena or as a "visible body action" (Kendon, 2010), movement criteria are effective measures to objectively describe hand movements. However, most gesture coding systems offer only vague descriptions of the movement form of a gesture type. Often it is not evident from the method described in a publication what kind of movement has actually been investigated. Hence, it is not possible to replicate the analysis and to confirm or disprove the findings, and thereby, to contribute to the body of knowledge concerning the validity of a particular movement value. Lott (1999) clearly demonstrated for research studies on the relationship between aphasia and gesture that the apparent contradictions between the results of different studies were essentially caused by the fact that the researchers had investigated movement values that, though they were termed similarly, actually referred to different movements. The imprecise definitions of the values entail insufficient objectivity and a lack of reliability. In fact, the examination of the interrater agreement is a somewhat recent development in the field of movement behavior and gesture research, and it is still not yet fully established as a standard method. Furthermore, many hand movement and gesture coding systems tend to operate with confounded values. For instance, gesture values such as *ideographics* (Efron, 1941) and *metaphorics* and *iconics* (McNeill, 1992) are confounded with linguistic assessments. These gesture types are primarily defined by the linguistic context and not with reference to the actual form of the movement per se. Confounded values do not enable the detection of gesture-speech mismatches, nor do they challenge existing paradigms such as the inseparability of gesture and speech production. Thus, the non-confounded registration of movement behavior is a prerequisite for investigating the validity of the movement values, as identified by their visual appearance with regard to personality traits, psychopathology, and cognitive, emotional, and interactive functions.

To summarize, coding systems for hand movements and gestures that ignore the bidirectional link between movement form and function and that rely on a paradigm in which each function can be realized by each movement form should be challenged. Furthermore, these coding systems, which operate with functionally defined values, are at risk of observers' assessments of the function or meaning of a hand movement being based on incorrect premises. The bidirectional link between movement function and movement form constitutes a solid paradigm for developing coding systems for hand movement behavior. Objective and unconfounded movement values that are defined by the visual appearance of the movement are a prerequisite for conducting basic research on the relationship between movement behavior and cognitive, emotional, and interactive functions.

The analysis of an ongoing stream of behavior

In hand movement behavior and gesture research, the ongoing stream of behavior is rarely submitted for analysis. Instead, a specific movement or gesture type, which is the focus of the respective research interests, is picked out of a stream of behavior. For instance, all pointing gestures are selected. This methodological approach is efficient with regards to the expenditure of time if only the analysis of specific types of movements is intended. However, it has the limitation that movements that do not at first sight seem to perfectly match the target prototype are neglected. Ambiguous forms or variations of a movement type, which may be identified only after exclusion of other types when analyzing the ongoing stream of behavior, might, however, provide valuable information about the movement type itself and about the associated cognitive, emotional, and interactive processes.

In contrast, if the ongoing stream of behavior is submitted for analysis, the researcher is forced to thoroughly consider each motion and to attribute a value to it. Thereby, the precision of the analysis and the gain in knowledge are substantially improved. Furthermore, hand movement and gesture analyses that are based on the segmentation and classification of an ongoing stream of movement behavior rather than on an a priori selection of certain movements provide a more reliable basis for quantitative analyses, since the variations of the target movement type are also registered.

The few researchers who have, thus far, analyzed an ongoing stream of behavior have employed different methods. Some researchers segment the behavior by units of time; for example, for each 30-s interval, they count the number of self-touches. This approach, however, destroys the natural units of behavior and it provides no information about the temporal structure of this type of behavior. For instance, a natural self-touching unit may last over 2 min. In that case, time unit coding would register four self-touch "units" instead of one unit the self-touch unit. Freedman and colleagues (1972) already pointed out that for some types of movement behavior such as *continuous body-focused* (self-touching) movements the number per time unit is not an effective measure.

The duration of natural units of a specific behavior constitutes an intrinsic feature of this behavior. If the mean unit duration of a specific movement value differs significantly from the mean unit duration of another value, this difference provides evidence that the two values represent behavioral entities that are distinct from each other. An example from NEUROGES is that *irregular on body* units last significantly longer than *phasic in space* units. Among others, this reflects the differences in the temporal extensions of the associated cognitive, emotional, and interactive processes of these two values. In the given example, the self-regulation processes associated with *irregular on body* movements last longer than the externalization of mental concepts associated with *phasic*

in space units. Therefore, the registration of the duration of the natural units of a specific movement value contributes to examining the validity of that value.

Furthermore, the analysis of the ongoing stream of movement behavior enables the understanding of its anatomy. When any motion is identified, the complete picture of hand movement behavior emerges and the relation between the different movement types becomes evident. Patterns can be detected—that is, recurrent sequences of different movement types in the course of time, or recurrent combinations of types displayed simultaneously by the right and left hands. In an intra-individually and intra-dyadically reliable manner, these movement patterns are associated with specific emotional, cognitive, and interactive states (e.g., Davis & Hadiks, 1994; Scheflen, 1973).

Laterality of hand movements

Many gesture coding systems do not provide values to analyze the relationship between the two hands (e.g., Efron, 1941; McNeill, 1992). However, lateral preferences for certain hand movement types—for example, a right-hand preference for pointing gestures or a left-hand preference for self-touch—indicate cerebral hemispheric specialization in the generation of the respective movement types. The anatomical basis for inferring hemispheric specialization from lateral preferences is that the left cerebral hemisphere controls the contralateral right limbs, and vice versa, the right cerebral hemisphere the contralateral left limbs. The relevance of this anatomical constellation for hand movement behavior and gesture research becomes most evident when examining split-brain patients (Gazzaniga, Bogen, & Sperry 1967; Lausberg et al., 2003; Lausberg, Davis, & Rothenhäusler, 2000; Lausberg, von Arnim, & Joraschky, 2007; McNeill, 1992; Sperry, 1968; Trope, Fishman, Gur, Sussman, & Gur, 1987; Volpe, Sidtis, Holtzman, Wilson, & Gazzaniga, 1982). These are patients in whom the corpus callosum, which is the biggest neural fiber connection between the right and left hemispheres, has been sectioned. As a result, distinct distal movements of the right and left hands can only be controlled by the contralateral left and right hemisphere, respectively. Thus, the left and right hand movements reflect competence or incompetence of the contralateral hemisphere regarding the generation of these movement types. Accordingly, neurologically healthy subjects tend to prefer the hand that is contralateral to the predominantly engaged hemisphere. As an example, Hampson and Kimura (1984) observed in right-handed healthy subjects a shift from right hand use in verbal tasks toward greater left-hand use in spatial tasks. According to their interpretation of the findings, the shift toward more left-hand use reflects the increased draw on particular right hemispheric competences during special tasks. Likewise, in behavioral

laterality experiments there is an advantage in responding with the hand that is controlled by the same hemisphere that performs the task (Zaidel, White, Sakurai, & Banks, 1988). This anatomical constellation contributes to explaining why in spontaneous unimanual hand movements, right-handers show a shift toward more left-hand use for self-touch, batons, and emotional gestures, whereas they prefer the right hand for pointing gestures and pictorial gestures that match the semantic content of their verbal utterances (Blonder, Burns, Bowers, Moore, & Heilman, 1995; Dalby, Gibson, Grossi, & Schneider, 1980; Foundas et al., 1995; Kimura, 1973a, b; Kita, de Condappa, & Mohr, 2007; Lavergne & Kimura, 1987; Saucier & Elias, 2001; Sousa-Poza, Rohrberg, & Mercure, 1979; Stephens, 1983; Trevarthen, 1996; Wilkins & de Ruiter, 1999). The registration of the laterality of hand movements and, furthermore, of the dominant hand in bilateral movements (Lausberg et al., 2000, 2007b) provides some indication of the hemisphere in which the movement type is generated. If a specific movement value is preferentially performed by the right hand or the left one, the preference suggests that this type is predominantly generated in the contralateral hemisphere. The hemispheric specialization for a certain movement type indicates that its generation is also associated with those cognitive and emotional processes that are also lateralized to that hemisphere.

Automated movement recognition

The current technical progress in automated movement recognition opens promising prospects for hand movement behavior and gesture research. However, thus far, in computer-based analysis, movement values are often chosen because they are simple to register. Although these values are objective and reliable, they often fail to be psychologically and functionally valid, as Krippendorf (180, p. 130) states it: "Reliability often gets in the way of validity." On the other hand, the values of traditional gesture coding systems (e.g., Efron, 1941; Ekman & Friesen, 1969; Kendon, 1990, 2010; McNeill, 1992) that are theoretically promising are typically not sufficiently operationalized to be suitable for automated approaches.

With reference to the aims and requirements outlined above, the first version of the NEUROGES–ELAN system was created, which is described in detail in the article by Lausberg and Sloetjes (2009) published in this journal. However, several aspects needed to be improved. Furthermore, new insights, partially based on experiences with the system in several empirical studies, led to changes in both the coding system and the annotation tool. In particular, the intention of the revision was to simplify the assessment and to avoid redundancy of assessments. A further goal was to achieve a more stringent conceptualization of the categories and values. Although these

changes were improvements in their own right, they should further improve interrater reliability and user-friendliness, render NEUROGES an even more suitable system for the development of automated movement recognition approaches, and facilitate the examination of the validity of the categories and values.

The following Method section describes the kinds of methodological shortcomings that were identified in the 2009 version of the NEUROGES–ELAN system and the strategies that were taken to address these shortcomings in the current revised version. Although some of these shortcomings might be specific to the NEUROGES–ELAN system, other deficits represent common problems in behavior research, and the methodological strategies taken to overcome them are of general interest for the improvement of behavior analysis. At the end of the Method section, a detailed overview on the most relevant revisions of the NEUROGES categories and values is provided. The direct comparison of the 2009 version and the current revised version enables users of the 2009 version to relate and discuss their findings with reference to studies using the current revised or interim versions.

The Results section starts with a survey of all empirical studies that have used the NEUROGES–ELAN system since its first publication in BRM in 2009. The survey illustrates the broad spectrum of fields of application of the system. Reliability as defined by Krippendorff (1980) was taken as a measure to assess whether the revisions had improved the NEUROGES–ELAN system. The improvements of interrater agreement scores in the seven NEUROGES categories reveal which methodological changes in the course of the revision process have been particularly effective. Finally, for each category the interrater agreement scores of the most recent studies, which have employed the present revised version, provide a frame of reference for interpreting agreement scores in future studies using the NEUROGES–ELAN system. Furthermore, these data are of interest because the EasyDIAg algorithm for assessing interrater agreement, which has recently been published in *Behavior Research Methods* (Holle & Rein, 2014) and has been applied consistently here across all of the studies analyzed, is new to the field, and thus far reference data have not been available.

Method

The revisions to the NEUROGES–ELAN system concerned the general structure of the NEUROGES system, single categories and values, as well as ELAN functions.

The revised structure of the NEUROGES system

The NEUROGES coding system is characterized by a vertical structure consisting of three subsequent modules that contain altogether seven subsequent categories (steps). It is a specific feature of the system that coding and segmentation constitute interdependent processes. In each category, on the basis of its specific movement criteria, a specific value is given to the movement unit. The unit is then adopted for the next coding step (category) and it is reassessed according to the specific movement criteria of that category. If in the next category the behavior changes within the adopted unit according to the criteria of that category, the unit is segmented into two (or more) subunits. These subunits then constitute the to-be-coded units for the next coding step. Because this principle of (sub)unit generation applies to all coding steps, the multistage evaluation process results in more and more fine-grained behavioral units. Thus, at Steps 6 and 7, when complex decisions concerning the function and meaning of the movement are required, fine behavioral units are available that are based on a highly operationalized step-wise segmentation of behavior.

The 2009 version emphasized the composition of the system with three modules: Module I: Kinetic hand movement coding, with the three steps (i) segmentation of behavior into movement units, (ii) trajectory and dynamics, and (iii) location of the action; Module II: Bimanual relation coding, with two steps (i) spatial relation, and (ii) functional relation; Module III: Functional gesture coding, with two steps: (i) function, and (ii) type. All modules were designed such that the value of the preceding step (category) determined the choice of values in the subsequent step. For instance, in Module I the Step 3 value *distant* could only be given to a unit if its Step 2 value was *phasic* or *repetitive*, or, in Module II the Step 2 value *symmetrical* only if the Step 1 value was *separate*.

The NEUROGES system is instantiated as an ELAN template file. In the 2009 template file, each module was represented by one Controlled Vocabulary (CV). The CV contained complex values that comprised all assessments conducted in one module. For instance, the data output of Module I were values such as *phasic on body*, which contained the Step 1 assessment (implicitly, by the fact that there is a movement unit), the Step 2 assessment (*phasic*), and the Step 3 assessment (*on body*). Accordingly, the interrater agreement was calculated for each module for these complex values (cf. Table 4 in the Results section).

In order to render the assessment more reliable, the system was first revised with the aim of clarifying the assessment process. Instead of documenting the rater's assessment after the final step of a module, it was documented after each assessment step. This was achieved by the introduction of the *Copy Tier* function in ELAN that allows the creation of a full duplicate of a tier (representing an assessment step and category, respectively), including its annotations (representing the values). For example, a *phasic* unit was transferred to the following tier (next category) and then was recoded with a complex value—for example, *phasic on body*, which contains

the value *phasic* in combination with the value of the new category. The CV for the modules was changed accordingly: It contained the simple values of the first step and the complex values that summarized the first and second steps. Accordingly, the interrater agreement scores were obtained for the simple and the complex values (cf. Table 5 in the Results section).

This procedure, however, still implied the redundant coding of the first step, and conceptually it reflected the dependency of a subsequent category on the preceding one. A stronger independency of the categories was strived for, because not only would this facilitate the localization of weaknesses in the conceptualization and operationalization of certain categories and values, but in the long run it would also promote the examination of validity. Hand in hand with the conceptual revision of the categories, the introduction of the *Concatenation* function in ELAN enabled the realization of this goal. The Concatenation function automatically generates a new tier with annotations based on the annotations of two input tiers. As an option, the values of overlapping annotations can be amalgamated in the new annotation. As an example, in the Structure category, a unit receives the value *phasic*, and in the independent Focus category, the copied unit receives the value *in space*. With the Concatenation function the two values can then be merged to the new value *phasic in space*. Thereby, the Concatenation function enables each step to attribute a value to a movement unit independently of the preceding step. Thus, in the most recent studies using the revised version of the NEUROGES–ELAN system, the interrater agreements refer to values that exclusively contain the assessment of one category (cf. Table 7 in the Results section). These revisions shifted the emphasis in the vertical structure of the NEUROGES system from a module-based approach to a category-based approach (see Fig. 1, Steps 1–7).

The separation of the steps methodologically disentangles the assessment process, and it also bears the advantage that the theoretical background behind each category can be better examined. The NEUROGES categories are each related to specific cognitive and neuropsychological functions (Lausberg, 2013). For instance, the Structure category reflects the complexity of the cognitive processes underlying the production of hand movements. The Focus category refers to the locus that attention processes are directed at. Within each category, the values are organized on a polar axis—for example, in the Focus category, from internal to external. The polar organization of the values, which had not yet been elaborated in the 2009 version, constitutes the new horizontal structure in the NEUROGES system (Fig. 1; see the order of the values within each step). The horizontal conceptualization implied the modification of values and the creation of new values (see below Table 1).

Another new feature in the structure of the NEUROGES system concerns the separation between (i) the segmentation of the ongoing stream of behavior and the analysis of all hand movements and (ii) the specific analysis of conceptual hand movements, which are identified by the preceding hand movement behavior analysis. In the 2009 version, basically all hand movements could be submitted to all assessment steps, from 1 through 7. This implied, for instance, that *phasic in space* movements (functionally these are gestures) and *shift* movements could be attributed the same Function value. However, considering the bidirectional link between form and function (as we argued in the Introduction), not all functions can be attributed to each movement form. Thus, it is not meaningful to use the same Function values for all types of hand movements. The revised version of the NEUROGES system takes into account that hand movements that differ in their Structure and Focus essentially rely on different cognitive, psychological, and neuropsychological processes. Therefore, in the revised version, only Steps 1–4 (Activation, Structure, Focus, and Contact categories) apply to all hand movements. The assessment Steps 5–7 (Formal Relation, Function, and Type categories) refer to more complex phenomena and are designed specifically for hand movements that are based on conceptual processes, such as gestures or tool use—that is, hand movements with a *phasic* or *repetitive* Structure (Fig. 1; see the horizontal bar between Steps 4 and 5).

Finally, redundancies in the assessment procedure were eliminated throughout the system. For instance, in Module I, as is described above, the recoding of the Structure values in the combined StructureFocus values was quit. In Module II, it was found that the information of the values *in touch* was already provided by the more fine-grained values of the subsequent step. Furthermore, the new Concatenation function in ELAN allowed us to drop the value *independent*, since it could be inferred from concatenation of the Structure values in the right and left hands. Thereby, in Module II the number of values could be reduced from 9 to 7. In Module III, the Type values of the Function values *object-oriented action* and *subject-oriented action* could be dropped, since the revision of the Focus category entailed that the information about the locus where the hands acted was already registered by the values *within body*, *on body*, *on attached object*, and *on separate object*.

Revisions of categories and values

In addition to the structural revisions described above, all categories have been revised conceptually. To facilitate the development of automated algorithms, in Module I the Structure category has been made more stringent with reference to the trajectory and to the phases that emerge from trajectory patterns. On the basis of new insights from empirical studies using the 2009 version, the three values of the Focus category have been differentiated into five more fine-grained values, and a new Focus value has been added (Table 1). In Module
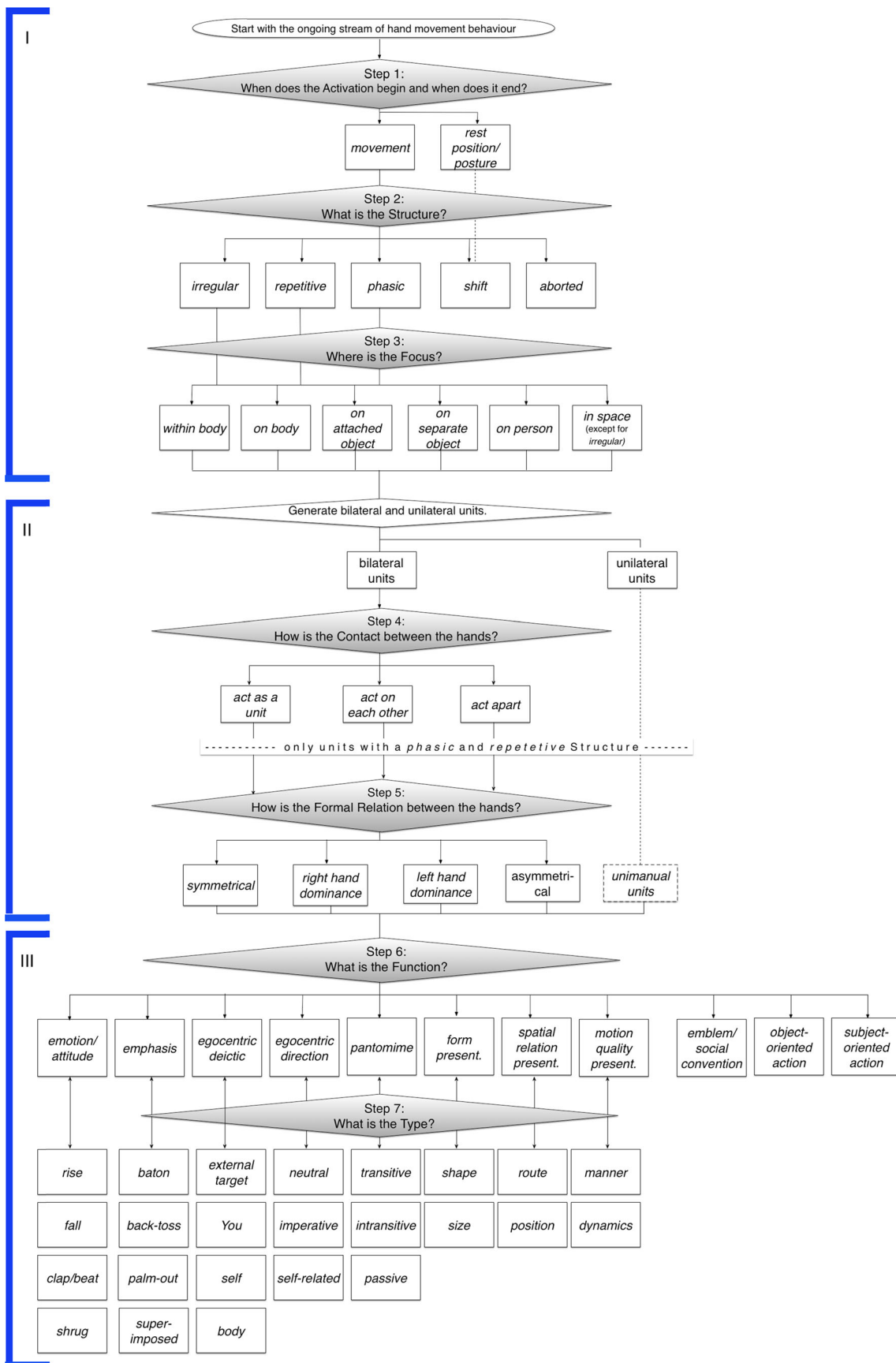
Start with the ongoing stream of hand movement behaviour

Step 1:
When does the Activation begin and when does it end?

| *movement* | *rest position/ posture* |

Step 2:
What is the Structure?

| *irregular* | *repetitive* | *phasic* | *shift* | *aborted* |

Step 3:
Where is the Focus?

| *within body* | *on body* | *on attached object* | *on separate object* | *on person* | *in space* (except for *irregular*) |

Generate bilateral and unilateral units.

| bilateral units | | unilateral units |

Step 4:
How is the Contact between the hands?

| *act as a unit* | *act on each other* | *act apart* |

- - - - - - - - - o n l y   u n i t s   w i t h   a   *p h a s i c*   a n d   *r e p e t e t i v e*   S t r u c t u r e - - - - - - -

Step 5:
How is the Formal Relation between the hands?

| *symmetrical* | *right hand dominance* | *left hand dominance* | *asymmetri- cal* | *unimanual units* |

Step 6:
What is the Function?

| *emotion/ attitude* | *emphasis* | *egocentric deictic* | *egocentric direction* | *pantomime* | *form present.* | *spatial relation present.* | *motion quality present.* | *emblem/ social convention* | *object- oriented action* | *subject- oriented action* |

Step 7:
What is the Type?

rise	*baton*	*external target*	*neutral*	*transitive*	*shape*	*route*	*manner*
fall	*back-toss*	*You*	*imperative*	*intransitive*	*size*	*position*	*dynamics*
clap/beat	*palm-out*	*self*	*self-related*	*passive*			
shrug	*super- imposed*	*body*					

Fig. 1 The revised version of the NEUROGES analysis system for hand movement behavior and gesture

Table 1 Comparison of the 2009 version and the revised version of the NEUROGES–ELAN system

2009 VERSION	REVISED VERSION
MODULE I	
Step 1	
(no category label)	**Activation***
movement	movement**
	rest position / posture
⁺ The new ELAN function *Create annotations from gaps* enables the automatic generation of *rest position/posture* units of the non-tagged time intervals between the *movement* units. Any moment in the continuous stream of movement behavior is thereby attributed to a (*movement* or *rest*) unit.	
Step 2	
Trajectory and Dynamics	**Structure**
In the revised definition of the category, the trajectory and the phases that emerge from trajectory patterns have been emphasized as the main criteria.	
phasic	phasic
repetitive	repetitive
continuous	irregular
The revised term *irregular* emphasizes the trajectory rather than the temporal dimension. Occasionally, units with an *irregular* trajectory are not continuous but are rather short in duration.	
shift	shift
stopped/holding	aborted
The revised term *aborted* prevents confusion with units that contain static complex phases ("stroke hold").	
Step 3	
Location of Acting	**Focus**
The category label Focus is compatible with the concept of the value *within body*.	
on body	on body
	on attached object
A pilot analysis on 34 males and 32 females revealed significant gender differences between direct acting on the body and acting on objects that are attached to the body. The latter behavior was displayed significantly more often by the females than by the males. Therefore, value *on body* from the 2009 publication was separated into *on body* and *on attached object*.	
distant	in space
	on separate object
The value *distant* was too imprecise, since in special study designs (e.g. (Helmich & Lausberg, 2014) participants manipulated the chin rest, the table, the arm rests, etc. These movements needed to be differentiated from *distant* movements such as gestures. Therefore, the 2009 value *distant* was separated into *in space* and *on separate object*.	
within body	within body
It was found that *within body* can co-occur with a *phasic* or *repetitive* Structure, such as repetitively rolling the shoulders. Therefore, in the revised version the Focus value *within body* is no longer restricted to *irregular* Structure units.	
	on person
For specific study designs such as studies of children (e.g. Bryjová et al., 2013) or on non-human primates closely interacting with each other, this value has been newly introduced.	
MODULE II	
	Generate bilateral and unilateral units
The new ELAN function *Annotations from overlap* enables the generation of bilateral units by creating units from overlapping right hand and left hand movement units. The new ELAN function *Annotations from subtraction* enables unilateral units to be generated. - The two Module II categories have been more clearly conceptually separated from each other and their values have been reorganized.	
Step 4	
Spatial Relation	**Contact**
The category label was changed to Contact in order to emphasize the physical contact between the two hands. - The word *act* was integrated in the value labels (and the label *folded* was deleted) in order to prevent confusions with the contact of the hands in *rest positions / postures*, e.g. resting with the hands folded.	
in touch	act on each other
	act as a unit
It was found that the 2009 values *folded/act as a unit, act on each other,* and *separate* actually refer to the type of physical contact between the two hands. Therefore, the three values constitute the new category Contact.	
separate	act apart
The term *act apart* is better than the label *separate* at indicating that the hands might act in a cooperative manner.	
	after this assessment step: **Only units with a *phasic* and *repetitive* Structure**
see above: revisions in the structure of the NEUROGES system	

Table 1 (continued)

Functional Relation	Formal Relation
Step 5	
Functional Relation	**Formal Relation**
The category label was changed from Functional to Formal in order to distinguish the category more clearly from the Function category in Module III.	
folded/act as a unit	see Contact category
act on each other	see Contact category
symmetrical	symmetrical
complementary	asymmetrical
The term *asymmetrical* was introduced for consistency due to the use of the term *symmetrical*.	
right hand dominance	right hand dominance
left hand dominance	left hand dominance
independent	
The value *independent* referred to the simultaneous display of a conceptual movement in one hand and a non-conceptual hand movement in the other hand. Because of the new Concatenation function that enables the merger of the Structure values of right hand and left hand units, this value has become redundant.	
MODULE III	
Step 6 Function / • **Step 7 Type*****	**Step 6 Function** / • **Step 7 Type**
emotion	emotion/attitude
The revised Function value *emotion/attitude* identifies hand movements that are exclusively a reflection of emotion and attitude (see above). With reference to movement direction and weight, four new Type values were defined.	
• intrinsic	• rise
	• fall
	• clap/beat
	• fall-shrug
emphasis	emphasis
• baton	• baton
	• superimposed
The value *superimposed* was introduced to distinguish *batons* that are added to another gesture type from primary *batons*.	
• back-toss	• back-toss
• palm-out	• palm-out
pointing	egocentric deictic
The term *egocentric* indicates that the revised Function value is restricted to those pointing gestures (deictics) that are displayed from an egocentric perspective. As reflected in his/her gesture, the gesturer uses her-/himself as the point of spatial reference. - All Type values of the Function value *egocentric deictic* have been labeled with respect to the target they are referring to.	
• deictic	• external target
	• you
The second person (You) was identified as a relevant target of pointing gestures.	
• self-deictic	• self
• body-deictic	• body
• hand-showing	
Since the Type value *hand-showing* occurred rarely, it was fused with the Type value *body*, which also refers to the gesturer's body.	
• direction	egocentric direction
It was found that the movement forms of gestures that indicate a direction or a route by using an egocentric frame of reference differ in an essential way from those indicating a location. Therefore, *direction,* which was only a Type value in the 2009 version, was reconceptualized as the new Function value *egocentric direction*. - With reference to the agent of the direction movement, three Type values are distinguished.	
	• neutral
	• imperative
	• self-related
motion	
In accordance with the differentiation between egocentric and (mento-) heliocentric perspectives (see above), the 2009 Function value *motion* was differentiated into *pantomime* (egocentric) and *motion quality* (heliocentric).	
	pantomime
• panto-image	
• panto-BPO	
The 2009 Type values *panto-image* and *panto-BPO* represented two modes of representation (demonstrating tool use with an imaginary object held in hand versus with the hand used as if it were the object) but were of the same Type. In the revised version, they are summarized in the Type value *transitive*, while the mode of representation is coded in a supplementary module.	
	• transitive
• panto-intransitive	• intransitive
	• passive
The value *passive* was added to register pantomime gestures in which an imaginary object acts on the gesturer, e.g. being hit by a stone.	
• kinetograph	see revised Function value *motion quality presentation*
• itinerary	see revised Function value *spatial relation presentation*

Table 1 (continued)

objects	*form presentation*
The label *form presentation* refers more clearly to the fact that the gesture primarily depicts the size and shape of a form (see new Type values). Thereby, it better distinguishes it from the Function value *pantomime,* which might include the depiction of objects.	
	• *shape*
	• *size*
• *ico-palpating*	
• *ico-BPO*	
see above, the modes of representation are now coded in a supplementary module	
space	*spatial relation presentation*
The revised Function value considers that in gesture, a spatial relation can only be recognized if at least two points in space are marked. - In contrast to the spatial Function values *egocentric deictic* and *egocentric direction,* the Function value *spatial relation presentation* is defined by a mento-heliocentric perspective. A *spatial relation presentation* that might include form information (see above, hierarchy).	
• *size/distance*	see Type value *size*
	see Type value *route*
The reference to one object as in the 2009 definition of the Type value *size/distance* was excluded. The *distance* aspect was included in the revised Type value *route.*	
• *trace*	• *route*
	see Type value shape
The 2009 Type value *trace* was compatible with a gesture showing a closed contour. Since this aspect primarily refers to an object, it is now integrated in the definition of the Type value *shape* (see above). Instead, the *distance* and *itinerary* aspects of the 2009 Type values (see above) are included in the revised Type value *route.*	
• *position/locus of reference*	• *position*
	motion quality presentation
In contrast to the motion Function value *pantomime,* the Function value *motion quality presentation* is defined by a mento-heliocentric perspective. With reference to the motion component, two Type values are distinguished.	
	• *manner*
	• *dynamics*
convention	*emblem / social convention*
Since emblems and social conventions differ between different cultures and subcultures, NEUROGES researchers are asked to set up their own list of *emblems / social conventions* for the culture they investigate. The NEUROGES coding manual provides a list of *emblems / social conventions* for studies on German speaking subjects.	
• *emblem*	see list
• *shrug*	see list
The 2009 Type value *shrug* is included in the *emblem* list. Excluded are shrugs that show a heavy fall of the shoulders (see above Type value *fall-shrug* of the Function value *emotion / attitude*).	
• *palming*	see list
Palming is included in the emblem list.	
• *rise-fall*	see Type values *rise* and *fall*
Since *rise-fall* gestures are mere expressions of emotion, in the revised version they are represented by Type values of the Function value *emotion / attitude.*	
objective purpose	*object-oriented action*
• *self-touch functional*	
• *body-adjacent manipulation functional*	
• *object manipulation functional*	
Since the locus where the hand acts on is already registered in the revised Focus category by the values *on body, on attached object,* and *on separate object,* the three Type values have become redundant.	
autostimulation	*subject-oriented action*
• *proprioception*	
• *self-touch autostimulating*	
• *body-adjacent autostimulating*	
• *object manipulation autostimulation*	
compare *object-oriented action*: The locus information is already registered by the revised Focus values *within body, on body, on attached object,* and *on separate object.*	
position shift	
• *(+ 2 Type values)*	
Since the revised Module III only assesses conceptual hand movements (with a *phasic* or *repetitive* Structure), units with a *shift* Structure are no longer considered.	
incomplete gestures	
• *(+ 4 Type values)*	
Since the revised Module III assesses conceptual hand movements based on to their complex phase, incomplete gestures consisting only of transport phases are no longer considered.	
several gestures in a unit	
This 2009 Function value implied a loss of information, since the single gestures were not classified. In the revised version, the segmentation of units into subunits is also possible in the Function and Type categories.	

*The category labels are marked by a capital letter. **The value labels are in *italics*.+The gray fields contain comments on the revisions. ***The Type category specifies the Function category; that is, almost each Function value is further specified by Type values.

II, the concepts of the two categories have been more clearly separated from each other. The new Contact category refers to the physical contact between the hands only, and the new Formal Relation category concentrates on the movement criteria of symmetry and dominance. In the Module III Function and Type categories, the values have been defined more stringently with respect to movement criteria. For example, the revised Function value *emotion/attitude* classifies hand movements with regard to trajectory, body involvement, laterality, movement direction, and the effort factor weight (based on embodiment theories and Laban movement analysis; Laban, 1988). Furthermore, supported by the findings from neurocognitive studies using the NEUROGES system (Sassenberg, Foth, Wartenburger, & van der Meer, 2011), the egocentric versus mento-heliocentric perspective reflected in the speaker's gestures has been operationalized, and it has been introduced as a new criterion to differentiate Function values. A further revision refers to *presentation* gestures, which may include various sorts of information. A form, a specific spatial relation, and a motion quality can all be conveyed in a single gesture. Therefore, a hierarchy between the different *presentation* Function values (*motion quality> spatial relation> form*) has been developed to define the relation between the three Function values. Table 1 provides an overview of the most relevant revisions of categories and values.

New developments in ELAN

ELAN has been under constant development, with two or three new releases per year. The main improvements developed in conjunction with the NEUROGES coding system concern those operations that create new annotations by combining the annotations of two (or more) input tiers while applying a logical operator (AND, OR, XOR). These are functions like Annotations from overlap, Merge annotations, Annotations by subtraction, Annotations from gaps, and so forth. The new option to concatenate the values of the annotations in these operations is one of the coordinated improvements of the NEUROGES–ELAN system. Further improvements concern multiple file processes (batch wise updating, converting or searching files) and calculation of interrater agreements (upcoming). The next version of ELAN will provide various algorithms for calculating interrater agreements; one of them is based on the EasyDIAg algorithm. Furthermore, the NEUROGES–ELAN system is designed in such a way that its output data can be directly submitted for statistical analyses. The codings are easy to export and convert into the variables required for statistical files. For SPSS users, NEUROGES .sav template files are available into which the NEUROGES output data can be inserted.

Table 2 Overview on empirical studies employing the NEUROGES–ELAN system

Grant or thesis project	Project name; study topic	Number and gender of participants; nationality; characteristics	Version of the NEUROGES system; coded steps	Author(s), year of publication; Study ID
2009				
Studienstiftung des Deutschen Volkes	Description of everyday activities	9 m, 18 f; German	2009: 2/3	Sassenberg & Helmich, 2013; H&S_13-3
Federal Ministry of Eduction and Research (BMBF) - NIL-I/II	Intelligence and brain structure	28 m; German	2009: 6-7	Wartenburger et al., 2010; W_10
Studienstiftung des Deutschen Volkes	Reasoning and gesture	37 m, 11 f; German	2009: 6-7	Sassenberg et al., 2011; S_11
Volkswagenstiftung II/82175	Doctor-patient interaction	1 m, 1 f; German; doctor, patient	2009: 6-7	Lausberg, 2011; L_11
German Research Association (DFG) LA 1249/1-3	Hemispheric specialisation: Wechsler Adult Intelligence (WAIS)	9 m, 7 f; Canadian, U.S. American; incl. callosotomy patients	2009: 2/3; 2013/revised: 4-5	Dvoretska & Lausberg, 2013; D&L_13
Volkswagenstiftung II/82175	Towards a grammar of gesture: Narration Mr. Bean	35 m, 23 f; German; incl. patients with left hemisphere and right hemisphere brain damage	2009: 2/3	Hogrefe et al., 2010; H_10
Diploma thesis	Dyadic interaction on free topics	20 m, 60 f; German	2009: 1	Dvoretska, 2009; D_09
Volkswagenstiftung II/82175	Towards a grammar of gesture: Narration Sylvester & Tweety	35 m, 23 f; German; incl. patients with left hemisphere and right hemisphere brain damage	2009: 1, 2/3; 2009: 4-5; 2013/revised: 6	Skomroch et al., 2013; Sk_13
2011 interim				
German Research Association (DFG) LA 1249/2-1	Hemispheric specialisation: Gestural demonstration of movement scenes	6 m, 5 f; Canadian	2011: 1-3; 2009/11: 4-5; 2013: 6	Helmich & Lausberg, 2014; H&L_14
German Research Association (DFG) LA 1249/2-1	Hemispheric specialisation: Narration of movement scenes	6 m, 5 f; Canadian	2011: 1-3; 2009/11: 4-5; 2013: 6	Helmich et al., 2014; H_14

Table 2 (continued)

Excellence Cluster Languages of Emotion	Languages of Emotion (LoE): Levels of Emotional Awareness	17 m; German	2011: 1-3	Sassenberg & Helmich, 2013; S&H_13-1
Excellence Cluster	Languages of Emotion (LoE): Hamburg Wechsler Adult Intelligence (HAWIE)	17 m; German	2011: 1-3	Sassenberg & Helmich, 2013; S&H_13-2
2013 interim				
Excellence Cluster Languages of Emotion	Languages of Emotion (LoE): Levels of Emotional Awareness	34 m, 32 f; German; incl. alexithymic	2013: 1-5, 7	Lausberg et al., 2015; L_15-1
Excellence Cluster	Languages of Emotion (LoE): Hamburg Wechsler Adult Intelligence (HAWIE)	34 m, 32 f; German; incl. alexithymic	2013: 1-5, 7	Lausberg et al., 2015; L_15-2
German Research Association (DFG) LA 1249/1-3	Hemispheric specialisation: Narration of Mouse	11 m, 9 f; Canadian, U.S. American; incl. callosotomy patients	2013: 1-5	Dvoretska & Lausberg, 2015; D&L_15
Dissertation project	Narration of Korean Dance and Ballet	30 m, 30 f; Korean, German	2013: 1-5, 7	Kim, in prep.; K_15
German Research Association (DFG) LA 1249/1-3	Stability pilot-study	3 m, 1 f; Slovak, Suisse	2013: 1-3	Lausberg, 2013; L_13
Dissertation project	SOPHONET: Interaction in Psychotherapy	5 m, 18 f; German; patients with sociophobia; 2 therapists (23 x)	2013: 1-3	Dvoretska et al., 2013; D_13
revised				
Volkswagenstiftung II/82175	Accuracy pilot study	1 f; Canadian	revised: 2	Rein, 2013; R_13
Federal Ministry of Eduction and Research (BMBF) 01UG1240D	AUVIS: Narration of Pear story in Papua New Guinea	33 m, 27 f; German, English, Papua Malay, Wooi	revised: 1-6	Skomroch et al., 2014; Sk_14
Swiss National Fonds (SNF) CRSII3_147673/1	Stress in early childhood	9; Suisse; children aged 3-5 years	revised: 1-3	Bryjovà et al., 2013; B_13
Master thesis project	KomPASS: Doctor - patient interaction	15 f, 1 m; German; 15 doctors, 1 simulation patient (15 x)	revised: 1-3	Gabor et al., 2014; G_14

The revisions of Module I resulted altogether in four versions (2009, published in this journal; 2011 interim, unpublished; 2013 interim, Lausberg, 2013; and the current revised version) that have all been applied in research studies. The revisions of Module II led to two versions (2009; 2013/revised [the revised version of Module II matches the 2013 version]), and those of Module III in three versions (2009; 2013 interim; and revised). In order to examine whether the revisions have improved the NEUROGES–ELAN system, a survey was conducted to gather all available studies that have employed the 2009 version, the interim versions, and the revised version. The studies were checked if they provided data on the stability, reproducibility, and accuracy (Krippendorff, 1980) of the NEUROGES–ELAN system.

Results

Survey on studies employing the NEUROGES–ELAN system

The survey included 18 studies from different scientific disciplines (psychology, neuropsychology, neurology, linguistics, and cultural anthropology) that had employed the NEUROGES–ELAN system since 2009 (Table 2). The studies were part of projects that dealt with the neurobiological correlates of different NEUROGES values; the association between hand movements and cognitive processes; the association between hand movements and emotional processes; the relationship between prosody, semantics, and hand movements in the segmentation of events in different cultures; hand movement behavior in binge eating disorders in childhood; stress and physiological and psychological well-being in early childhood; and hand movement behavior in psychotherapy. The studies included 467 participants altogether, 182 of whom took part in two or three different studies. The participants were from different cultures (Germans, US Americans, francophone and anglophone Canadians, Swiss, Koreans, and Papua New Guineans), including healthy individuals as well as individuals with brain damage or mental illness.

The studies were examined as to whether they provided information concerning the reliability of the NEUROGES system. Following a recommendation by Krippendorff (1980), three types of reliability were distinguished: (i) stability as measured by intra-observer test–retest conditions, (ii) reproducibility as measured by inter-observer test-test conditions, and (iii) accuracy as measured in test–standard conditions. With the exception of two pilot studies, all studies provided data on reproducibility. In these studies, each participant's videotaped hand movement behavior was coded without sound by at least two independent trained raters. The two

pilot studies were designed explicitly to examine system stability and accuracy.

Cohen's kappa was only applied to Module III in the early studies that used the 2009 version. In all other studies, the EasyDIAg algorithm for assessing interrater agreement (Holle & Rein, 2014) was applied for Steps 2–7. The EasyDIAg score not only takes into account the raters' agreement about values (as Cohen's kappa) but also the raters' agreement about the segmentation of the behavior—that is, if the raters agree on when a unit starts and ends, and when the next unit begins. In contrast to the classical Cohen's kappa, which calculates the interrater agreement for a category, EasyDIAg provides an interrater agreement score for each value of a category. The EasyDIAg algorithm is described in detail in Holle and Rein (2014). Since EasyDIAg cannot be applied to binary categories, for the Activation category—the only binary one in the NEUROGES system—the procedure proposed by Petermann, Skomroch, and Dvoretska (2013) was used to calculate interrater agreement. It is based on the ELAN functions *overlap* and *merge*, and it enables the generation of a ratio between the total length of the overlaps in *movement* units of both raters and the total length of *movement* units of both raters.

The consistent application of EasyDIAg and the overlap–merge ratio scores since 2009 allows the comparison of the reproducibility of the 2009 version, the interim versions, and the revised version of the NEUROGES system.

Module I

Activation category Since there were no changes in the definition of the *movement* unit between the 2009 version and the

Table 3   Activation category: Overlap–merge ratio scores of studies the NEUROGES–ELAN system

Authors, Year of Publication	Study ID	Merge-Overlap Ratio ($M \pm SD$)
Dvoretska, 2009	D_09	.74± .11
Skomroch et al., 2013	Sk_13	.81 ± .16
Helmich & Lausberg, 2014a, b	H&L_14	.89± .23
Helmich et al., 2014	H_14	.82± .09
Lausberg et al., 2015 (Study 1[*])	L_15-1	rh:[**] .71± .30, lh: .74± .31
Lausberg et al., 2015 (Study 2[*])	L_15-2	rh: .74± .29, lh: .81± .23
Dvoretska & Lausberg, 2015	D&L_15	rh: .77± .13, lh: .73± .14
Kim, 2015	K_15	rh:.86± .14, lh: .85± .14
Dvoretska et al., 2013	D_13	rh: .87± .12, lh: .83± .16
Skomroch et al., 2014	Sk_14	rh: .84± .14, lh: .78± .19
Bryjovà et al., 2013	B_13	rh: .77± .03, lh: .78± .01
all studies (**M± SD**)		**.80± .05**

[*] Studies 1 and 2 ($n = 66$) included the Studies 1 and 2 by Sassenberg & Helmich (2013, $n = 17$); [**] rh = right hand; lh = left hand

Table 4 StructureFocus categories: EasyDIAg mean scores and raw agreement scores (in brackets) in studies using the 2009 version of the NEUROGES–ELAN system

Study ID	phasic distant	phasic on body	repetitive distant	repetitive on body	continuous* on body	continuous* within body	shift	stopped/ holding**
S&H_13-3	.35	.35	.51	.32	.50	.80	.43	.42
D&L_13	.67 (.91)	.48 (.89)	.55 (.94)	.62 (.94)	.59 (.90)	.00 (1.0)	.61 (.85)	.00 (1.0)
H_10	.61 (.82)	.50 (.90)	.67 (.93)	.53 (.96)	.71 (.94)	.00 (1.0)	.51 (.96)	.60 (.99)
Sk_13	.46 (.79)	.35 (.94)	.57 (.90)	.40 (.96)	.44 (.87)	.12 (.95)	.34 (.91)	.30 (.92)
all studies (M± SD)	**.52± .15**	**.42± .08**	**.58± .07**	**.47± .13**	**.41± .30**	**.23± .39**	**.47± .12**	**.33± .25**

* *Continuous in the 2009 version was re-conceptualized as irregular in the 2011 interim version;* ** *Stopped/holding in the 2009 version was retermed aborted in the 2011 interim version.*

revised version of NEUROGES, the interrater agreements from all of the studies since 2009 are listed in Table 3. In the most recent studies, the Activation scores were calculated separately for the right and left hands.

Structure and Focus categories Since 2009, the Structure and Focus categories have been revised three times, resulting in the 2011 interim version, the 2013 interim version, and the current revised version.

In the 2009 version of the NEUROGES–ELAN system, raters directly coded the combined StructureFocus values. Thus, the interrater agreement EasyDIAg scores were only calculated for the combined StructureFocus values (Table 4).

The mean ($M± SD$) of the six StructureFocus values was .44± .12, that of the two Structure values *shift* and *stopped/holding* was .40± .10, and that of all eight values was .43± .11.

Table 5 Structure and StructureFocus categories: EasyDIAg mean scores and raw agreement scores (in brackets) in studies using the 2011 interim version of Module I

Study ID	phasic	repetitive	irregular	shift	aborted
H&L_14	.88 (.96)	.90 (.98)	.66 (.98)	.59 (.93)	.39 (.95)
H_14	.70 (.93)	.78 (.98)	.67 (.91)	.61 (.90)	.62 (.96)
S&H_13-1	.63	.75	.46	.45	.56
S&H_13-2	.55	.71	.16	.26	.50
M ± SD	**.70 ± .14**	**.79 ± .08**	**.49 ± .24**	**.48 ± .16**	**.52 ± .10**

phasic distant**	phasic on body *	repetitive distant**	repetitive on body *	irregular-on body *	irregular within bod.
.88 (.96)	.54 (.99)	.92 (.98)	.57 (.99)	.68 (.98)	.50 (.99)
.72 (.94)	.66 (.99)	.70 (.98)	.80 (1.0)	.75 (1.0)	.56 (.96)
.64	.56	.72	.74	.50	.37
.59	.49	.71	.72	.30	.33
.71 ± .13	**.56 ± .07**	**.76 ± .11**	**.71 ± .10**	**.56 ± .20**	**.44 ± .11**

Values that conceptually match 2009 values are in the shaded columns. Note that in the 2011 interim version the value *on body* still included *on attached object* and the value *distant* still included *on separate object* and *in space* (in the publications H&L_14 and H_14 the term *in space* was used instead of *distant*, since *on separate object* occurred very infrequently).

In the 2011 interim version, the Structure category was assessed separately, but the three Focus values were assessed only in combination with the Structure values (Table 5).

The mean EasyDIAg score ($M± SD$) of the Structure values was .60± .14, and that of the StructureFocus values was .62± .12. The comparison of the 2009 and 2011 scores shows an improvement in all shared values (shaded columns)—that is, the six StructureFocus values and the two Structure values *shift* and *aborted*. Independent-samples t tests revealed that the increases in the EasyDIAg scores reached significance for *phasic on body* [$t(6) = –2.636, p = .039$], *repetitive distant* [$t(6) = –2.990, p = .024$], and *repetitive on body* [$t(6)= –2.901, p= .027$].

In the 2013 revision of the NEUROGES–ELAN system, the Focus category was elaborated. The number of Focus values increased from three to six, resulting in 17 combined StructureFocus values. The Structure values remained unchanged. The 2013 version was introduced during the course of two large-scale studies (L_15-1, L_15-2), when 17 of 66 subjects had already been assessed with the 2009 version of the StructureFocus category (S&H_13-1, S&H_13-1). These 17 participants were then reassessed with the 2013 version. Thus, the 2009 version and the elaborated 2013 version of the StructureFocus category could be directly compared. Table 6 shows the EasyDIAg scores from the Structure category and the elaborated StructureFocus category of the 2013 version.

The mean ($M± SD$) of the Structure values was= .59± .05, and that of the StructureFocus values was .65± .20. Six StructureFocus values in the 2013 version were conceptually comparable to values in the 2009/2011 versions (shaded columns). For all six values, independent-samples t tests showed significant improvements of the EasyDIAg scores: *phasic in space* versus *phasic distant* [$t(11)= 2.995, p= .012$]; *phasic on body* (version 2013) versus *phasic on body* (version 2009/11) [$t(11) = 4.147, p = .002$]; *repetitive in space* versus *repetitive distant* [$t(11) = 4.027, p = .002$]; *repetitive on body* (version 2013) versus *repetitive on body* (version 2009/11) [$t(11) = 3.619, p = .004$]; *irregular on body* (version 2013) versus *continuous/irregular on body* (version 2009/11) [$t(11) = 2.676, p = .022$]; and *irregular within body* (2013) versus

Table 6 Structure and StructureFocus categories: EasyDIAg mean scores and raw agreement scores (in brackets) for the Structure values in studies using the 2013 interim version of Module I

Study ID	phasic	repetitive	irregular	shift	aborted
D&L_15	.41 (.82)	.51 (.91)	.49 (.77)	.52 (.86)	.28 (.97)
L_15-1	.54 (.81)	.60 (.94)	.60 (.81)	.59 (.95)	.44 (.99)
L_15-2	.62 (.88)	.78 (.98)	.56 (.79)	.66 (.91)	.37 (.99)
K_15	.66 (.85)	.76 (.92)	.67 (.89)	.63 (.96)	.79 (.99)
D_13	.57 (.87)	.62 (.96)	.61 (.93)	.80 (.97)	.70 (.99)
L_13**	-	-	-	-	-
M ± SD	**.56 ± .10**	**.65 ± .11**	**.59 ± .07**	**.64 ± .10**	**.52 ± .22**

phasic on body*	phasic attached object	phasic separate object	phasic within body	phasic in space	repetitive on body	repetitive attached object
.70 (.96)	.64 (.99)	.83 (1.0)	.00 (1.0)	.82 (.95)	.91 (.99)	1.0 (1.0)
.78 (.99)	.64 (.99)	.00 (1.0)	.21 (1.0)	.93 (.97)	.91 (1.0)	.00 (1.0)
.81 (.98)	.62 (.99)	.30 (.99)	.89 (1.0)	.87 (.96)	.92 (1.0)	.75 (1.0)
.66 (.94)	.81 (1.0)	.54 (1.0)	.57 (1.0)	.82 (.93)	.75 (.97)	.40 (1.0)
.63 (.97)	.54 (.98)	.74 (.97)	.65 (.98)	.77 (.96)	.90 (1.0)	.60 (1.0)
.94	1.0	1.0	-***	.96	.79	-
.72 ± .08	**.65 ± .10**	**.48 ± .34**	**.46 ± .36**	**.84 ± .06**	**.88 ± .07**	**.55 ± .38**

repetitive separate object	repetitive within body	repetitive in space	irregular on body	irregular attached object	irregular separate object	irregular within body
-	-	.93 (.99)	.93 (.97)	.00 (1.0)	.75 (.99)	.86 (.97)
.00 (1.0)	-	.95 (.99)	.84 (.93)	.46 (.98)	.49 (.99)	.69 (.93)
.50 (1.0)	-	.87 (.99)	.86 (.93)	.34 (.98)	.48 (.98)	.71 (.94)
0	1.00	.91 (.98)	.75 (.93)	.47 (.99)	.57 (1.0)	.53 (.99)
.60 (1.0)	.50 (1.0)	.89 (.99)	.64 (.96)	.00 (1.0)	.99 (1.0)	.52 (.99)
-	-	.95	1.0	1.0	-	-
.28 ± .32	**.75 ± .35**	**.91 ± .03**	**.80 ± .11**	**.32 ± .22**	**.70 ± .14**	**.70 ± .14**

Values that conceptually match 2011 values are in the shaded columns. * The study designs in the cited studies did not allow for direct bodily interaction. Therefore, StructureFocus values with the Focus value *on person* did not occur. ** This was a pilot study on stability in which the same rater coded the same videos after a time interval of 2 years. The stability study is not included in the mean ± *SD* (last row) that applies only to reproducibility studies. *** The value did not occur in the study.

continuous/irregular with body (version 2009/11) [$t(11) = 2.376, p = .037$].

The current revised version of the NEUROGES–ELAN system is characterized by the fact that the Focus values are coded independently of the Structure values. Thus, they are no longer coded in combination with StructureFocus values. Furthermore, conceptually the Structure values are defined more strictly with regard to trajectory. Table 7 shows the EasyDIAg scores of the Structure and Focus categories in the revised version.

The EasyDIAg scores of all Structure values from the 2015 version ($M \pm SD = .78 \pm .07$) improved in comparison to the

Table 7 Structure category and Focus category: EasyDIAg mean scores and raw agreement scores (in brackets) in studies using the current revised version of the NEUROGES–ELAN system

Study ID	phasic	repetitive	irregular	shift	aborted
Sk_14	.63 (.82)	.68 (.90)	.75 (.90)	.68 (.97)	.38 (.97)
B_13	.96 (.99)	.94 (.99)	.92 (.98)	.95 (.99)	.92 (.99)
G_14	.73 (.91)	.74 (.97)	.84 (.92)	.85 (.99)	.71 (1.0)
M ± SD	**.77 ± .17**	**.79 ± .14**	**.84 ± .09**	**.83 ± .14**	**.67 ± .27**

within body	on body	on attached object	on separate object	on person	in space
.64 (.96)	.68 (.87)	.60 (.98)	.73 (.98)	-	.81 (.91)
.41 (.99)	.91 (.98)	1.0 (1.0)	.98 (1.0)	1.0 (1.0)	.89 (.99)
.82 (.96)	.91 (.95)	1.0 (1.0)	.00 (1.0)	-	.98 (.99)
.62 ± .21	**.83 ± .13**	**.87 ± .23**	**.57 ± .23**	**1.00**	**.89 ± .08**

Values that conceptually match 2011/13 values are in the shaded columns.

scores from the 2011/13 versions (.59 ± .10). In independent-samples t tests, these improvements were significant for *shift* units [$t(10) = 2.635, p = .025$] and for *irregular* units [$t(10) = 2.946, p = .01$]. Since the separate Focus values (.80 ± .17) were new, no comparison was possible with the previous versions.

In addition to these studies on reproducibility, one pilot study dealt with the accuracy of the Structure values (Rein, 2013). The five Structure values *phasic*, *repetitive*, *shift*, *aborted*, and *irregular* were reliably distinguished by 3-D kinematography. The kinematographic classification matched the human raters' assessments.

Module II

Module II has only been revised once, in 2013. That version is still valid and part of the current revised version of the NEUROGES system. Three studies are available that used the 2009 version of Module II (Table 8), and six studies have applied the 2013/revised version of Module II (Table 9).

Since in the revision of Module II some values were shifted between the two categories, for comparison purposes the EasyDIAg mean scores are reported for both categories together. As compared to the 2009 version's Spatial and Functional Relation values ($M \pm SD = .72 \pm .21$), the combined EasyDIAg score for the 2013/revised version's Contact and Formal Relation values (.74 ± .07) improved slightly (Contact category, .77 ± .10; Formal Relation, .72 ± .05). The independent-samples t tests were not significant for those seven values that were conceptually comparable between the two versions (shaded columns).

Module III

The Function category of Module III has been revised once. Its 2013 version is still valid and part of current revised

Table 8 Spatial Relation and Functional Relation categories: EasyDIAg scores from studies using the 2009 version of the NEUROGES–ELAN system

Study ID	in touch	separate	complementary	independent	left-hand dominance	right-hand dominance	symmetrical	folded/act as a unit	act on each other
H&L_14	.93 (1.0)	.95 (.98)	.61 (.97)	.01 (.98)	1.0 (1.0)	.87 (.99)	1.0 (1.0)	.92 (1.0)	0.73 (0.99)
H_14	.88 (.97)	.85 (.95)	.39 (.97)	.27 (.98)	.80 (1.0)	.71 (.95)	.65 (.94)	.32 (.98)	.94 (.99)
Sk_13	.79 (.96)	.88 (.95)	.59 (.99)	.55 (.97)	.76 (.97)	.77 (.96)	.87 (.97)	.51 (.99)	.82 (.97)
M±SD	**.87± .07**	**.89 ±. 05**	**.53± .12**	**.28± .27**	**.85± .13**	**.78 ±.08**	**.84± .18**	**.58± .31**	**.83± .11**

version. The Type category has been revised twice, resulting in a 2013 interim version and the current revised version. Since in the NEUROGES analysis the Type category is always the last step to be assessed in a project, no data on the revised Type category are currently available. However, most of the 2013 interim Type values match those of the current revised version.

The first studies that used the 2009 version of Module III still applied the classical Cohen's kappa (Cohen, 1960; Table 10). This was acceptable, since in Module III the segmentation of behavioral units into subunits plays a minor role, as compared to the previous assessment steps. Thus, the raters' disagreement mainly refers to categorial rather than temporal assessments.

Since the Type category is a specification of the Function category, researchers have decided either to code the Function category (four studies) or to code the Type category (two studies) in order to render the analysis less time-consuming. Tables 11 and 12 show the EasyDIAg scores for the Function and Type categories, respectively.

The mean score ($M± SD$) of the Function values was .62± .13, and that of the Type values was .69± .23. In order to compare the reproducibility of the Function values with the reproducibility of the Type values, for each group of Type

values that specified a Function value, the mean EasyDIAg score was calculated (see the columns labeled "Type mean"). The mean Function and Type EasyDIAg scores were compared via independent-samples t tests. With the exception of the values *emotion*, *egocentric deictic*, and *pantomime*, the mean scores of the Type values that specified a Function value were higher than the scores of the corresponding Function values. However, no results in the independent-samples t tests were significant, apart from a trend for *motion quality presentation* [$t(3.414)= –2.804$, $p= .058$].

Discussion

The aim of the present study was to examine the reliability of the system and, in particular, whether the revisions improved reproducibility.

The study survey yielded 18 empirical studies that have used the NEUROGES–ELAN system since 2009. Sixteen of them fulfilled the design requirements for the examination of reproducibility (Krippendorff, 1980). In these studies, the subjects' hand movement and gestural behavior were analyzed by at least two independent raters (video analysis without sound). In two pilot studies, we examined stability and accuracy, respectively.

In the following paragraphs for each of the NEUROGES modules, the data on reproducibility are discussed, and if possible, compared between the 2009 version, the interim versions, and the revised version.

Table 9 Contact and Formal Relation categories: EasyDIAg scores from studies using the 2013/revised version of Module II

Study ID	act as a unit	act on each other	act apart	asymmetrical	left hand dom.	right hand dom.	symmetrical
D&L_13	.59 (.94)	.85 (.95)	.77 (.87)	.46 (.95)	.52 (.98)	.63 (.96)	.88 (.95)
L_15-1	.89 (.99)	.96 (.98)	.95 (.98)	.62 (.93)	.77 (.96)	.73 (.95)	.77 (.90)
L_15-2	.70 (.98)	.91 (.96)	.93 (.96)	.83 (.94)	.81 (.96)	.80 (.95)	.85 (.93)
D&L_15	.41 (.96)	.81 (.91)	.77 (.89)	.44 (.98)	.62 (.99)	.64 (.97)	.74 (.97)
K_15	.59 (.99)	.66 (.92)	.72 (.87)	.50 (.89)	.76 (.97)	.79 (.97)	.68 (.87)
Sk_14	.68 (.99)	.85 (.95)	.70 (.87)	.61 (.85)	.70 (.95)	.65 (.92)	.82 (.92)
Mean ± SD	**.66 ± .16**	**.85 ± .83**	**.79 ± .14**	**.68 ±. 15**	**.70 ± .11**	**.71 ± .08**	**.79 ± .07**

Values that conceptually match 2009 values are in the shaded columns.

Table 10 Function and Type categories: Cohen's Kappa scores of studies using the 2009 version of the NEUROGES–ELAN system

Study ID	Cohen's Kappa
W_10	only bimanual gestures were assessed: κ= .75
S_11	right hand gestures κ= .69, left hand gestures κ= .67, bimanual gestures κ= .75
L_11	right hand, left hand, and bimanual gestures κ= .82

Table 11 Function category: EasyDIAg scores of studies using the 2013/revised version of the Function category

Study ID	emotion/ attitude	emphasis	egocentric deictic	egocentric direction	pantomime	form presentation	spatial relation presentation	motion quality presentation	emblem	object-oriented action	subject-oriented action
Sk_13	.69 (.99)	.86 (.95)	.59 (.99)	.56 (.99)	.81 (.96)	.68 (.97)	.58 (.97)	.75 (.93)	.75 (1.0)	.00 (1.0)	.87 (.96)
H&L_14	.38 (.94)	.72 (.96)	.00 (.98)	1.0 (1.0)	.67 (.84)	.61 (.95)	.00 (.99)	.43 (.85)	-*	-*	.54 (.93)
H_14	.88 (.99)	.70 (.92)	1.0 (1.0)	.39 (.98)	.73 (.92)	.32 (.97)	.49 (.98)	.47 (.49)	-*	-*	.66 (.84)
Sk_14	.70 (.98)	.73 (.91)	.64 (.97)	.53 (.94)	.74 (.95)	.65 (.97)	.63 (.92)	.71 (.96)	.83 (.99)	.71 (.99)	.77 (.96)
M±SD	**.66± .21**	**.75± .07**	**.56± .41**	**.62± .26**	**.74± .06**	**.57± .16**	**.43± .29**	**.59± .16**	**.79± .06**	**.36± .50**	**.71± .14**

* The study designs did not elicit *emblems* or *object-oriented actions*.

Module I

Module I comprises the Activation, Structure, and Focus categories.

The Activation category had not been revised since 2009. Eleven empirical studies provided data on the interrater agreement for the Activation category, as measured by an overlap–merge ratio. The mean ratio was .80, ranging from .71 to .89. Thus, 4/5 of the time raters agreed on whether there was *movement* or *rest*. Although at first glance this decision seems to be trivial, it turned out to be the most difficult measure to achieve interrater agreement on. This is mainly due to the fact that a *movement* unit cannot simply be operationalized by motion in space. For instance, a gesture might contain a static phase in which the hand is just held in space against gravity—for example, when displaying the Peace sign by holding the hand with V-shaped fingers. In this example, the motionless phase constitutes an intrinsic phase of a *movement* unit. In contrast, there may be motions of the hand that are not *movement* units. For instance, if the hand rests on the knee and the knee is shifted. In this example, the hand is (passively) in motion but it does not display a *movement*. These and many more pitfalls are all described in the NEUROGES coding manual, explaining the complications in the assessment *movement* versus *rest*, which constitutes the basic step in the segmentation of the continuous stream of behavior. In the current development of the automated algorithm of the NEUROGES system (Schreer, Masneri, Lausberg, & Skomroch, 2014), likewise, the Activation category was the most difficult step (O. Schreer, 2015, personal communication; S. Masneri, 2015, personal communication). Against this background, the mean overlap–merge ratio of .80 can be considered to be a good interrater agreement.

Since 2009, the Structure and Focus categories in Module I underwent three revisions. Four studies were conducted using the 2009 version, four with the 2011 interim version, five with the 2013 version, and three, thus far, with the revised version presented in this article.

In the 2009 version, only the combined StructureFocus category was assessed. The EasyDIAg scores for these complex values ranged between .23 and .58. Although these scores appear to be numerically low, a pilot study by Skomroch (2013) has related the EasyDIAg scores to classical Cohen's kappa scores and provides a frame of reference for their interpretation. Skomroch first calculated the EasyDIAg scores for the Structure category, and then he submitted the annotation data into a filter that removed the effects of the raters' disagreement about segmentation. Thus, as in classical Cohen's kappa, only the raters' categorial agreement created the agreement score. There was a substantial increase in the numerical scores for all five Structure values (*phasic* .46 → .85, *repetitive* .56 → .85, *irregular* .38 → .79, *shift* .35 → .75, and *aborted* .36 → .84). Although the filter procedure is not perfect, it nevertheless provides an impression of the impact of the raters' disagreement concerning the segmentation on the EasyDIAg scores.

The first revision of the Structure and Focus categories (2011) aimed at simplifying the coding process. The Structure category was assessed separately, followed by the assessment of the combined StructureFocus category. As compared to the 2009 version, the EasyDIAg scores for all StructureFocus assessments were improved, significantly for the values *phasic on body*, *repetitive distant*, and *repetitive on body.* Thus, the separate assessment of the Structure category before the StructureFocus assessment proved to be a compelling revision to improve the reproducibility of the StructureFocus values.

The next revision (2013) implied a conceptually motivated differentiation of the Focus category into six Focus values, resulting in 17 combined StructureFocus values. The revision led to a significant improvement in the eight StructureFocus values that could be compared with those that had existed in the preceding versions. The scores for the new StructureFocus values were clearly lower, ranging from .28 to .75. These new values occurred only infrequently and in some studies not at all, a fact that was likely to negatively influence the raters' recognition of these values.

This problem was solved in the current revised version, in which the assessment of a high number of complex StructureFocus values was replaced by the complete separation of the assessments of the Structure and Focus categories. A further conceptual change was the stricter definition of

Table 12 Type category (Type values specifying the Function values): EasyDIAg scores of studies using the 2013 version of the Type category

emotion/attitude

Study ID	rise	fall	clap/beat	shrug	palming	fist clenching	opening	closing	Type mean
L_15-1	.71 (1.0)	.68 (.95)	-	.81 (.97)	.00 (1.0)	1.0 (1.0)	.00 (1.0)	.44 (1.0)	.52
L_15-2	.76 (1.0)	.80 (.98)	-	.75 (.98)	-	-	.61 (1.0)	.82 (1.0)	.75
M±SD	.74±.04	.74±.08	-	.78±.04	.00	1.0	.31±.43	.63±.27	.74±.01

emphasis

Study ID	baton	super-imposed	back-toss	palm-out	Type mean
L_15-1	.80 (.99)	.72 (.99)	0.66 (0.97)	.88 (.96)	.77
L_15-2	.91 (1.0)	.66 (1.0)	0.79 (0.99)	.88 (.98)	.81
M±SD	.86±.08	.69±.04	.73±.09	.88±.00	.79±.03

egocentric deictic

Study ID	self	you	external target	body	Type mean
L_15-1	.78 (1.00)	.60 (1.00)	.38 (1.0)	-	.59
L_15-2	.00 (.99)	.59 (.99)	.00 (.99)	-	.20
M±SD	.39±.55	.60±.01	.19±.27	-	.59±.00

pantomime

Study ID	transitive	intransitive	Type mean
L_15-1	-	.67 (1.0)	.67
L_15-2	0.18 (0.99)	.67 (1.0)	.43
M±SD	.18	.67±.00	.55±.17

form presentation

Study ID	shape	size	Type mean
L_15-1	1.0 (1.0)	.75 (1.0)	.88
L_15-2	.80 (1.0)	.57 (1.0)	.69
M±SD	.90±.14	.66±.13	.79±.13

egocentric direction

Study ID	neutral	imperative	self-related	Type mean
L_15-1	.83 (1.0)	.80 (1.0)	.89 (.98)	.84
L_15-2	.69 (1.0)	.77 (1.0)	.86 (.99)	.77
M±SD	.76±.10	.79±.02	.88±.02	.81±.05

spatial relation presentation

Study ID	route	position	Type mean
L_15-1	.87 (1.0)	.82 (.99)	.85
L_15-2	.77 (1.0)	.90 (1.0)	.84
M±SD	.82±.07	.86±.06	.84±.01

motion quality presentation

Study ID	manner	dynamics	Type mean
L_15-1	.79 (.99)	.82 (.99)	
L_15-2	.84 (1.0)	.86 (1.0)	
M±SD	.82±.04	.84±.03	.83±.03

Study ID	emblem/soc. convention	object-oriented action	subject-oriented action
L_15-1	-	-	.87 (.95)
L_15-2	-	-	.87 (.94)
M±SD	-	—	.87±.00

The Type values are ordered according to the Function values that they specify.

Structure values with reference to the trajectory and trajectory patterns. This revision led to an improvement in the EasyDIAg scores for all Structure values. The improvement was significant for the Structure values *irregular* and *shift*. Furthermore, a kinematographic pilot study indicated the accuracy of the five Structure values. Since the six Focus values were assessed separately for the first time, no comparison with the previous NEUROGES versions was possible. The mean EasyDIAg score of the separate Focus values was .80 (.57–.89). Overall, the last revision, which was motivated conceptually and methodically by the aim of breaking down the assessment process into simple steps, was highly effective.

To summarize, all three revisions of the Structure and Focus categories, no matter whether they were conceptually or methodologically motivated, have significantly improved the reproducibility of the two categories.

Module II

Module II was only revised once. In that revision, the two categories were conceptually more strictly separated with regard to movement criteria. Furthermore, the redundancy in the assessments of the two categories was eliminated. The current version has existed since 2013.

The 2009 version already showed relatively good EasyDIAg scores (M = .72). The revision resulted in the exclusion of the value with the worst agreement (*independent*). Furthermore, the reproducibility of the second worst value (*complementary*) was improved in its revised counterpart (*asymmetrical*). Otherwise, for the seven conceptually shared values, the comparison of the EasyDIAg scores in the 2009 version and the 2013/revised version yielded no significant differences.

Module III

In Module III, the Function category was revised once. The current Function category version has existed since 2013. The Type category was revised twice, resulting in the 2013 version and in the current revised version.

In the 2009 version of Module III, the classical Cohen's kappa was used. The kappa scores for the Function and Type categories in the three studies ranged between .69 and .82. According to the classification scheme for Cohen's kappa proposed by Landis and Koch (1977), these scores indicate substantial to almost perfect agreement. The more conservative scheme by Shrout (1998) would indicate moderate to substantial agreement. Thus, the 2009 version of Module III already demonstrated good reproducibility.

The revisions of Module III were mainly conceptually motivated. In particular, they aimed at a stricter operationalization of the Function and Type values based on

movement criteria. The Function category refers to broader groups of movements that share certain movement features that are associated with specific cognitive, emotional, and interactive functions. In contrast, the Type category defines smaller groups of movements that share even more movement features. In order to answer the methodological question of whether the reproducibility would be better for the broader Function values or for the more fine-grained Type values, researchers were asked to assess either the Function category (in four studies) or the Type category (in two studies). The EasyDIAg mean score of the 11 Function values was .62 and that of the 24 Type values .69. The higher mean score for the Type values suggests that the raters profited more from the fine-grained behavioral values. The comparison of the eight Function values that are specified with Type values and the means of the corresponding Function-specifying groups of Type values showed that for *emphasis*, *egocentric direction*, *form presentation*, *spatial relation presentation*, and *motion quality presentation*, better scores were achieved in the Type category. Thus, when assessing *presentation* gestures in particular, raters seemed to profit from fine-grained values. However, more research studies are needed to verify this assumption.

To summarize, the EasyDIAg scores for the revised NEUROGES categories range between .62 and .80. Since EasyDIAg is a new measure for reproducibility, thus far, no classification schemes for the interpretation of the scores are available. As we indicated above, standard classification schemes for Cohen's kappa scores cannot be applied to EasyDIAg scores. Indirect evidence for the strength of agreement is provided by Module III of the 2009 version in which the Cohen's kappa scores showed substantial to almost perfect and moderate to substantial, respectively, agreement. Furthermore, the segmentation filter procedure by Skomroch (2013) allows relating the EasyDIAg scores to Cohen's kappa classification schemes. According to this procedure, the Structure values showed at least the same strength of reproducibility as the Module III values. Clearly, future research should aim at developing a classification scheme for EasyDIAg scores.

In the current revised version of the NEUROGES–ELAN system, the EasyDIAg scores were best for Module I (.78–.80), followed by Module II (.72–.77), and then Module III (.62–.69). The differences between the modules might reflect on one hand the amount of conceptual and methodological elaboration invested in the development of a module and on the other hand the complexity of the behavioral phenomena submitted for analysis. In contrast, the segmentation of behavior, which is the most frequently required in the first steps of the revised NEUROGES–ELAN analysis and decreases from Step 1 to Step 7, seems to contribute less to the raters' disagreement. More research will be needed to clarify for each category how different factors influence the EasyDIAg score.

The present study survey provides a frame of reference for comparing EasyDIAg scores of a category between different studies. As an example, in Modules I and II the scores in the study by Helmich and Lausberg (2014; study ID H&L_14) were consistently above average. A factor that explains the good reproducibility in that study is the highly structured design in which gesture production without speech was investigated in standardized stimulus–response conditions. Finally, the present study survey enables the assessment of the quality of the different NEUROGES–ELAN versions relative to each other. It has clearly shown that for all modules, the revisions have effectively improved reproducibility.

In line with the present confirmation of the reliability of the NEUROGES system, the study survey showed that, indeed, the tool has been used for basic empirical research. As a descriptive and comprehensive system with unfounded values based on the visual appearance of the movement only, the NEUROGES–ELAN system has been applied in the investigation of a variety of topics on hand movement behavior and gesture in relation to personality, level of intelligence, state of mental health, quality of interaction, effectiveness of psychotherapy, cognitive processes, brain anatomy, and hemispheric specialization.

In accordance with the original aims of the tool NEUROGES–ELAN system has been used as an interdisciplinary tool. Furthermore, the two-dimensional structure of the NEUROGES system implies that each module and each category provide specific results that are valid per se—that is, independent of the findings in other categories. The survey indicated that researchers profited from the flexible structure of NEUROGES. In six studies, only Module I was used; in three studies, only Module III; in two studies, Modules I and II; and in seven studies, all Modules I–III (Step 6 or Step 7). The vertical structure further entails that NEUROGES can be combined with existing coding systems for nonverbal behavior. Thus, its flexible structure might, indeed, have promoted the use of NEUROGES–ELAN across scientific disciplines.

Conclusion

The original aim for the development of NEUROGES–ELAN was to create an objective and reliable tool for the analysis of an ongoing stream of hand movement behavior. The system should enable basic research on hand movement behavior and gesture in relation to cognitive, emotional, and interactive functions. Furthermore, it should be comprehensive and sensitive to complex behavioral phenomena, but remain user-friendly and suitable for automated recognition approaches. To serve an interdisciplinary community of researchers, the system should be flexible in its use and compatible with existing coding systems.

In line with these original goals, a revision of the 2009 version of the NEUROGES–ELAN was conducted. The assessment procedure was simplified by avoiding redundant assessments, and clarified by replacing complex assessment steps with simple ones. Categories and values were conceptualized in ways that were more stringent with reference to movement criteria.

The present survey of 18 empirical studies demonstrates that the revisions have substantially improved the reproducibility of results. The revised NEUROGES–ELAN system has proven to be a reliable system for the analysis of hand movement behavior and gesture, and its use across scientific disciplines characterizes it as an effective tool for interdisciplinary research.

References

Allport, G. W., & Vernon, P. E. (1933). *Studies in expressive movement*. New York, NY: Macmillan.

Blonder, L. X., Burns, A. F., Bowers, D., Moore, R. W., & Heilman, K. M. (1995). Spontaneous gestures following right hemisphere infarct. *Neuropsychologia, 33,* 203–213.

Bryjovà, J., Munsch, S., Puder, J., & Hoffmann, J. (2013). *Die Reaktivität auf einen akuten Stressor und Emotionsregulationsmechanismen bei Kindern im Vorschulalter*. Paper presented at the Workshop Kongress für Klinische Psychologie und Psychotherapie, Trier, Germany.

Burn, H. (1987). The movement behavior of anorectics—The control issue. *American Journal of Dance Therapy, 10,* 54–76.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Dalby, J. T., Gibson, D., Grossi, V., & Schneider, R. D. (1980). Lateralized hand gesture during speech. *Journal of Motor Behavior, 12,* 292–297.

Davis, M., & Hadiks, D. (1994). Nonverbal aspects of therapist attunement. *Journal of Clinical Psychology, 50,* 393–405.

Dvoretska, D. (2009). *Kinetische Interaktion und Koordination*. Diploma thesis, Humboldt-Universität, Berlin.

Dvoretska, D., & Lausberg, H. (2013). *Verbal and nonverbal intelligence show different patterns of hand movement*.Paper presented at the 55th Conference of Experimental Psychologist, Vienna, Austria.

Dvoretska, D., Joraschky, P., & Lausberg, H. (2013). *Kinesische Interaktion als Prädiktor für Therapieerfolg*. Poster presented at the Deutscher Kongress für Psychosomatische Medizin und Psychotherapie, Heidelberg, Germany.

Dvoretska, D., & Lausberg, H. (2015). *How we change our gestural behavior in different situations?* Paper presented at the SPSP Pre-Conference on Nonverbal Behavior, Long Beach, California.

Efron, D. (1941). *Gesture and culture*.]The Netherlands: Mouton.

Eisenberg, P. (1937). Expressive movement related to feelings of dominance. *Archives of Psychology, 211*.

Eisenberg, P., & Reichline, P. B. (1939). Judgements of dominance feelings from motion pictures of gait. *Journal of Abnormal and Social Psychology, 10,* 345–357.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage and coding. *Semiotica, 1,* 49–99.

Fahrenberg, J., Hampel, R., & Selg, H. (1984). *Das Freiburger Persönlichkeitsinventar (revidierte Fassung FPI-R und teilweise geänderte Fassung FPI-A1* (4th ed.). Göttingen, Germany: Hogrefe.

Foundas, A. L., Macauley, B. L., Raymer, A. M., Mahler, L. M., Heilman, K. M., & Rothi, L. J. G. (1995). Gesture laterality in aphasic and apraxic stroke patients. *Brain and Cognition, 29,* 204–213.

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 153–175). New York, NY: Pergamon.

Freedman, N., O'Hanlon, J., Oltman, P., & Witkin, H. A. (1972). The imprint of psychological differentiation on kinetic behavior in varying communicative contexts. *Journal of Abnormal Psychology, 79*(3), 239–258.

Frijda, N. H. (1965). Mimik und Pantomimik. In R. Kirchhoff (Ed.), *Handbuch der Psychologie, Band 5, Ausdruckspsychologie* (pp. 351–422). Göttingen, Germany: Hogrefe.

Gabor, N., Vitinius, F., Obliers, R., Ritter, A., & Lausberg, H. (2014). *Gestural behavior as an indicator for measuring the quality of a therapist–patient relationship.* Poster presented at the International Conference on Interactional Competences in Institutional Practices (ICIP), Paris, France.

Gazzaniga, M. S., Bogen, J. E., & Sperry, R. W. (1967). Dyspraxia following division of the cerebral commissures. *Archives of Neurology, 16,* 606–612.

Goldenberg, G. (1996). Defective imitation of gestures in patients with damage in the left or right hemispheres. *Journal of Neurology, Neurosurgery, and Psychiatry, 61,* 176–180.

Goldenberg, G. (1999). Matching and imitation of hand and finger postures in patients with damage in the left or right hemispheres. *Neuropsychologia, 37,* 559–566.

Goldenberg, G., Laimgruber, K., & Hermsdörfer, J. (2001). Imitation of gestures by disconnected hemispheres. *Neuropsychologia, 39,* 1432–1443.

Goldenberg, G., & Strauss, S. (2002). Hemispheric asymmetries for imitation of novel gestures. *Neurology, 59,* 893–897.

Hampson, E., & Kimura, D. (1984). Hand movement asymmetries during verbal and nonverbal tasks. *Canadian Journal of Psychology, 38,* 102–125.

Hargadine, M. (1973). Development and criticism of a measurement instrument for scope of movement. Paper presented at the Eighth Annual Conference of the American Dance Therapy Association, Overland Park, Kansas.

Helmich, I., & Lausberg, H. (2014). Hand movements with a phase structure and gestures that depict action stem from a left hemispheric system of conceptualization. *Experimental Brain Research, 232,* 3159–3173. doi:10.1007/s00221-014-4006-x

Helmich, I., Skomroch, H., & Lausberg, H. (2014). Neuropsychological functions of hand movements and gestures change in the presence or absence of speech. *Journal of Cognitive Psychology, 26,* 740–753.

Hogrefe, K., Skomroch, H., Rein, R., Goldenberg, G., & Lausberg, H. (2010). *Gesture production in the right and left hemispheres during narration of short movies.* Paper presented at the Gesture—Evolution, Brain, and Linguistic Structures, 4th Conference of the International Society for Gesture Studies (ISGS), Frankfurt am Main, Germany.

Holle, H., & Rein, R. (2014). EasyDIAg: A tool for easy determination of interrater agreement. *Behavior Research Methods. Advance online publication.*. doi:10.3758/s13428-014-0506-7

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language, 89,* 253–260.

Kendon, A. (1990). *Conduction Interaction. Patterns of behavior in focused encounters.* New York, NY: Cambridge University Press.

Kendon, A. (2010). *Gesture—visible action as utterance* (4th ed.). Cambridge, UK: Cambridge University Press.

Kim, Z.H. (2015). Vergleich des verbalen Handbewegungsverhaltens zwischen Deutschen und Koreanern zur Messung ihrer Wahrnehmung von unterschiedlichen Tanzstimuli des Balletts und koreanischen Tanzes "Sung-Mu" [Comparison of verbal gesture behavior between German and Korean as an indicator for measuring their perception of different dance stimuli of ballet and Korean dance "Sung-Mu"], German Sport University, doctoral thesis in preparation.

Kimura, D. (1973a). Manual activity during speaking—I. *Right-handers. Neuropsychologia, 11,* 45–50.

Kimura, D. (1973b). Manual activity during speaking—II. *Left-handers. Neuropsychologia, 11,* 51–55.

Kimura, D., & Archibald, Y. (1974). Motor functions of the left hemisphere. *Brain, 97,* 337–350.

Kita, S., de Condappa, O., & Mohr, C. (2007). Metaphor explanation attenuates the right-hand preference for depictive co-speech gestures that imitate actions. *Brain and Language, 101,* 185–197. doi:10.1016/j.bandl.2006.11.006

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology.* Newbury Park, CA: Sage.

Laban, R. (1988). *The Mastery of Movement,* 4th ed. rev. and enl. Plymouth, UK: Northcote House.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Lausberg, H. (1998). Bewegungsdiagnosetest mit Bewertungsskalen für Diagnostik und Therapieevaluation in der Tanztherapie. *Zeitschrift für Tanztherapie, 7,* 35–42.

Lausberg, H. (2011). Das Gespräch zwischen Arzt und Patientin: Die bewegungsanalytische Perspektive. *Balint Journal, 1,* 15–24.

Lausberg, H. (Ed.). (2013). *Understanding body movement: A guide to empirical research on nonverbal behavior. With an introduction to the NEUROGES coding system.* Frankfurt/Main, Germany: Peter Lang.

Lausberg, H., & Cruz, R. F. (2004). Hemispheric specialization for imitation of hand-head positions and finger configurations: a controlled study in patients with complete callosotomy. *Neuropsychologia, 42,* 320–334.

Lausberg, H., Cruz, R. F., Kita, S., Zaidel, E., & Ptito, A. (2003). Pantomime to visual presentation of objects: Left hand dyspraxia in patients with complete callosotomy. *Brain, 126,* 343–360.

Lausberg, H., Davis, M., & Rothenhäusler, A. (2000). Hemispheric specialization in spontaneous gesticulation in a patient with callosal disonnection. *Neuropsychologia, 38,* 1654–1663.

Lausberg, H., Helmich, I., Sassenberg, U., & Petermann, K. (2015). *Gender-specific implicit emotional processing in alexithymic individuals.* Paper presented at Deutscher Kongress für Psychosomatische Medizin und Psychotherapie, Berlin, Germany.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES–ELAN system. *Behavior Research Methods, 41,* 841–849. doi:10.3758/BRM.41.3.841

Lausberg, H., von Arnim, A., & Joraschky, P. (2007a). Bewegungsanalyse in der Körperbild- und Körperschemadiagnostik. In G. H. P. Geißler (Ed.), *Lebensbewegungen: Zum körperlichen Geschehen in der psychoanalytischen Therapie. Ein Lehrbuch* (pp. 165–196). Heidelberg, Germany: Springer.

Lausberg, H., von Wietersheim, J., & Feiereis, H. (1996). Movement behavior of patients with eating disorders and inflammatory bowel disease: A controlled study. *Psychotherapy and Psychosomatics, 65,* 272–276.

Lausberg, H., Zaidel, E., Cruz, R. F., & Ptito, A. (2007b). Speech-independent production of communicative gestures: Evidence from

patients with complete callosal disconnection. *Neuropsychologia, 45,* 3092–3104. doi:10.1016/j.neuropsychologia.2007.05.010

Lavergne, J., & Kimura, D. (1987). Hand movement asymmetry during speech: No effect of speaking topic. *Neuropsychologia, 25,* 689–693.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). Oxford, UK: Oxford University Press.

Lott, P. (1999). *Gesture and aphasia.* Bern, Switzerland: Peter Lang.

Mason, D. J. (1957). Judgments of leadership based upon physiognomic cues. *Journal of Abnormal and Social Psychology, 54,* 273–274. doi:10.1037/h0042054

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago, IL: University of Chicago Press.

Petermann, K., Skomroch, H., & Dvoretska, D. (2013). Calculating temporal interrater agreement for binary movement categories. In H. Lausberg (Ed.), *Understanding body movement: A guide to empirical research on nonverbal behavior. With an introduction to the NEUROGES coding system* (pp. 253–260). Frankfurt/Main, Germany: Peter Lang.

Reid, D. W., & Ware, E. E. (1974). Multidimensionality of internal-external locus of control: Implications for past and future research. *Canadian Journal of Behavioral Science, 3,* 264–271.

Rein, R. (2013). Using 3D kinematics for the analysis of hand movement behavior: A pilot study and some further suggestions. In H. Lausberg (Ed.), *Understanding body movement: A guide to empirical research on nonverbal behavior. With an introduction to the NEUROGES coding system* (pp. 187–190). Frankfurt/Main, Germany: Peter Lang.

Rimoldi, H. J. A. (1951). Personal tempo. *Journal of Abnormal and Social Psychology, 46,* 283–303. doi:10.1037/h0057479

Sassenberg, U., Foth, M., Wartenburger, I., & van der Meer, E. (2011). Show your hands—Are you really clever? Reasoning, gesture production, and intelligence. *Linguistics, 49,* 105–134.

Sassenberg, U., & Helmich, I. (2013). Statistical evaluation and data presentation. In H. Lausberg (Ed.), *Understanding body movement: A guide to empirical research on nonverbal behavior. With an introduction to the NEUROGES coding system* (pp. 289–322). Frankfurt/Main, Germany: Peter Lang.

Saucier, D. M., & Elias, L. J. (2001). Lateral and sex differences in manual gesture during conversation. *Laterality, 6,* 239–245.

Scheflen, A. E. (1973). *Communicational structure: Analysis of a psychotherapy transaction.* Bloomington, IN: Indiana University Press.

Schreer, O., Masneri, S., Lausberg, H., & Skomroch, H. (2014). *Coding hand movement behavior and gesture with NEUROGES supported by automated video analysis.* Paper presented at the 9th International Conference on Methods and Techniques in Behavioral Research (Measuring Behavior 2014), Wageningen, The Netherlands.

Shostrom, E. L. (1972). *EITS manual for the personal orientation inventory.* San Diego, CA: Educational and Industrial Testing Service.

Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods for Medical Research, 7,* 301–317.

Skomroch, H. (2013). Recommendations for assessing the level of interrater agreement. In H. Lausberg (Ed.), *Understanding body movement: A guide to empirical research on nonverbal behavior. With an introduction to the NEUROGES coding system* (pp. 279–286). Frankfurt/Main, Germany: Peter Lang.

Skomroch, H., Lausberg, H., Unterladstetter, V., Sandler, M., Strunk, J., & Himmelmann, N. (2014). *Patterns of hand movement and prosodic behavior across language groups.* Paper presented at the Mapping Multimodal Dialogue 2 Conference, Leuven, Belgium. https://www.arts.kuleuven.be/midi/events/mamud2-program.pdf

Skomroch, H., Petermann, K., Helmich, I., Dvoretska, D., Rein, R., Kim, Z.H., … Lausberg, H. (2013). *Gender differences in hand movement behavior.* Poster presented at the Tilburg Gesture Research Meeting (TiGeR), Tilburg, The Netherlands. Retrieved from tiger.uvt.nl/pdf/papers/skomroch.pdf

Sousa-Poza, J.F., Rohrberg, R., & Mercure, A. (1979). Effects of type of information (abstract–concrete) and field dependence on asymmetry of hand movements during speech. *Perceptual and Motor Skills, 48*(3, Part 2), 1323–1330.

Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23,* 723–733. doi:10.1037/h0026839

Stephens, D. (1983). *Hemispheric language dominance and gesture hand preference.* PhD dissertation: Department of Behavioral Sciences, University of Chicago, Chicago, Illinois.

Trevarthen, C. (1996). Lateral asymmetries in Infancy: Implications for the development of the hemispheres. *Neuroscience & Biobehavioral Reviews, 20,* 571–586.

Trope, I., Fishman, B., Gur, R. C., Sussman, N. M., & Gur, R. E. (1987). Contralateral and ipsilateral control of fingers following callosotomy. *Neuropsychologia, 25,* 287–291.

Volpe, B. T., Sidtis, J. J., Holtzman, J. D., Wilson, D. H., & Gazzaniga, M. S. (1982). Cortical mechanisms involved in praxis: Observations following partial and complete section of the corpus callosum in man. *Neurology, 32,* 645–650.

Wallbott, H. G. (1989). Movement quality changes in psychopathological disorders. In B. Kirkaldy (Ed.), *Normalities and abnormalities in human movement* (pp. 128–146). Basel, Germany: Karger.

Wartenburger, I., Kühn, E., Sassenberg, U., Foth, M., Franz, E. A., & van der Meer, E. (2010). On the relationship between fluid intelligence, gesture production, and brain structure. *Intelligence, 38,* 193–201.

Wilkins, D., & de Ruiter, J. P. (1999). Hand preference for representational gestures: A comparison of Arente and Dutch speakers. In V. Van Geenhoven & N. Warner (Eds.), *Annual Report 1999* (pp. 51–52). Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.

Zaidel, E., White, H., Sakurai, E., & Banks, W. (1988). Hemispheric locus of lexical congruity effects: Neuropsychological reinterpretation of psycholinguistic results. In C. Chiarello (Ed.), *Right hemisphere contributions to lexical semantics* (pp. 71–88). New York, NY: Springer.