RESPONSE

Grip Formation as an Emergent Property. Response to Commentaries on "A New View on Grasping"

Jeroen B.J. Smeets and Eli Brenner

We begin our response by discussing the commentators' arguments concerning our proposal to abandon the classical distinction between transport and grip. In the second section, we argue that the minimum-jerk model is not fundamental to our approach, but very convenient. In the third section, we discuss how the experimental results that the commentators mention fit into our new approach. We conclude that the predictive capacity of our model, combined with its simplicity, makes it very useful for understanding grasping.

Let us begin our response by citing the man who initiated recent grasping research: Marc Jeannerod. In a recent paper written with some colleagues (Jeannerod et al., 1995), he states, "The question of why grip aperture is larger than that required by object size is still a matter of debate" (p. 314). We have formulated a new answer to this question in our target article: it is the consequence of the general strategy to approach surfaces more or less perpendicularly. None of the commentators disputes this strategy. However, neither the wrist nor the grip closure is directed toward a point on a surface. On the other hand, each digit approaches a point on a surface, and shows the tendency to do this perpendicularly. If the strategy is indeed general, we must therefore assume that grip formation emerges from more or less independent movements of the digits. At this point the views start to diverge.

Abandoning Visuomotor Channels for Transport and Grip

Most commentators support our step to abandon the classical description of transport and grip. Newell (we cite only the corresponding author when referring to commentaries written by more than one author) argues that we are still trapped by Jeannerod's line of thought. In part this is true; for instance, we adhere to the concept of visuomotor channels. In other aspects, however, our line of thought may only appear similar to that of Jeannerod because we discuss the predictions of our model in terms of the classical variables. We used these variables in order to compare our predictions with the published experimental data.

Both Steenbergen and Savelsbergh present additional support for our view that the anatomical argument for the classical view is not very strong. In Section

The authors are with Vakgroep Fysiologie, Erasmus Universiteit Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands.

1.2 we discussed that the direct corticospinal projections to intrinsic muscles of the hand appear not to be crucial for the reach to grasp, but are important *after* the digits have made contact with the object (Lemon et al., 1995). This role of the corticospinal projections is supported by recent findings by Steenbergen et al. (1998), who studied the problems in grasping that children with cerebral palsy have. They found that these problems occur *after* contact has been made with the object. Savelsbergh argues that the same relationship between the transport and grip component is observed in various tasks with quite different effectors. This also clearly raises doubts about any anatomical argument in favor of the classical view. We will return to this issue when discussing experimental predictions.

Marteniuk argues that the classical view has withstood 18 years of research and should therefore not be abandoned too readily. In his interpretation, our theoretical objections against the classical view are not very strong. However, the aspect in Jeannerod's classical description that we theoretically dispute is not the (independent) control of transport and grip, but that these parameters of motor behavior are linked directly to perceptual information in independent visuomotor channels. This very attractive hypothesis is still supported by Jeannerod and co-workers (see for instance Paulignan & Jeannerod, 1996). By discussing the various aspects of the relationship between various object properties and grip orientation, we showed (Section 1.1) that it is impossible to consistently define the classical two independent visuomotor channels. In his commentary, Rosenbaum gives an additional argument as to why information about object position is not enough to guide the wrist; information on finger aperture is needed as well.

As an alternative to the classical assumption of independent visuomotor channels for transport and grip, we assumed separate visuomotor channels for finger and thumb. Rosenbaum and Savelsbergh question the proposal that finger and thumb are controlled independently. For this argument, they cite work by Cole and colleagues (Cole & Abbs, 1986; Cole et al., 1984) in which subjects were asked to pinch. They found that a perturbation of the movement of one digit resulted in a strong response in the other. Pinching, however, differs in one important aspect from grasping. In pinching, the task is to move the thumb and the finger to each other, instead of to a position defined in external space. A perturbation of the movement of one digit thus automatically changes the target position for the other digit. In our view, such a change in target position should result in a response from the other digit. Our claim (which we still have to test) is that when such a perturbation is applied while grasping an object, the trajectory of the unperturbed digit is unaffected.

Savelsbergh opts for the approach proposed by Wing (thumb and grip are controlled) as an alternative for the classical description. Although the work of Wing and colleagues inspired us, their approach neglects the problems in relating grip aperture to the size of the object. As we discussed in Section 1.1, the size of the grip not only depends on the size of the object but also on the positions on the object at which the digits will make contact. Thus, the final positions of both finger and thumb have to be planned. Moreover, we have shown experimentally that the grip aperture is not determined by the apparent size of the object (Brenner & Smeets, 1996).

Stelmach argues that Jeannerod's original model is outdated because in some situations its components are interdependent, and in others they are not. To incorporate this finding, he proposes to extend the classical description with a higher

order control system to coordinate the transport and grip component. In our opin-
ion, introducing an extra coordinator is not very elegant. Moreover, it is not neces-
sary if one assumes that transport and grip are not controlled, but instead are emer-
gent properties of the simultaneous control of the digits' movements. We modeled
the control of these movements and conclude from our calculations that in simple
symmetric situations, our model behaves exactly as if there were independent con-
trol of transport and grip. However, if we introduce an asymmetry in the model
situation (e.g., treating finger and thumb slightly differently, as argued in Section
3.1 of the target article), the emerging transport and grip are both related to intrin-
sic object properties (i.e., the value of a_p, see Figure 1). According to our model,
the (symmetry in the) task determines whether transport and grip appear indepen-
dent, without any need to introduce changes in the digits' control.

One of the main advantages of our approach is that we do not need a sepa-
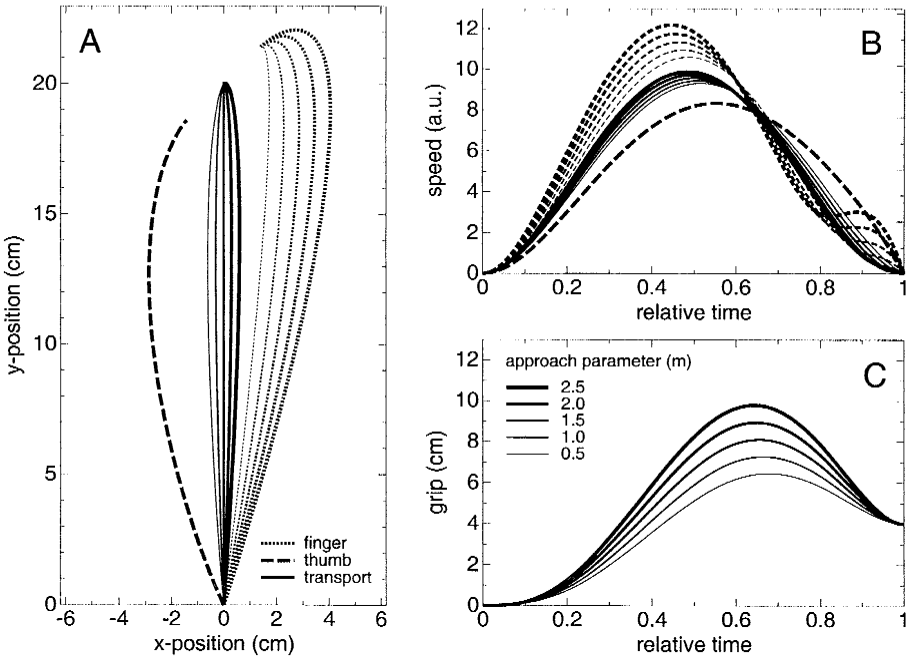rate model for grasping. Any model that describes pointing toward a surface will

Figure 1 — A set of trajectories generated by our model for various values of the approach
parameter (a_p) for the finger, and a constant a_p for the thumb. Disks 4 cm in diameter are grasped
at 20-cm distance. The a_p for the finger is 0.5 to 2.5 m, and for the thumb it is 1.5 m. The thickness
of the curves is proportional to the a_p of the finger. (A): Calculated paths of finger and thumb and
their average (transport component). Path of transport component depends on the a_p for the
finger. (B): Velocity profiles of the movements of finger and thumb and of the resulting transport
component. Shape of transport velocity profile depends on the a_p for the finger. In this geometry,
peak velocity occurs earlier if the a_p for the finger is larger. (C): Time-course of grip aperture as
derived from the calculated trajectories of the digits.

describe grasping as well. Steenbergen argues that grasping and pointing are controlled differently, because grasping has a clear purpose whereas pointing does not. In our view, many pointing tasks have a clear purpose. While writing this reaction, my fingers point in sequence to various keys to press them far enough to result in words on my monitor. It has been shown recently (Brenner & Smeets, 1995; Klein Breteler et al., 1998) that when pointing movements are made toward real objects, movements are curved in a way that is compatible with a perpendicular approach.

To support his claim that grasping and pointing are controlled differently, Steenbergen cites a study (Carnahan et al., 1993) in which responses to changes in target position were examined. In that study, the response of the transport component of a grasping movement was compared with the response of a finger during pointing. It was found that the perturbation reduced the time to peak velocity from 210 to 190 ms in grasping, whereas it remained the same (180 ms) in pointing. These results are not conclusive, because they are compatible with the hypothesis that the response has a latency of about 190 ms in both tasks. However, even if there is a real difference, it does not necessarily contradict our hypothesis that the control of movements of the *digits* is comparable in both tasks. Moreover, the digits' movements will only be the same if the constraints to be met at the point of contact are the same. In the study by Carnahan et al. (1993), the pointing and grasping tasks differ in several respects. Subjects probably pointed less accurately than they grasped, as revealed by the shorter movement times. Moreover, the perturbation was perpendicular to the contact surface (the dowel) in grasping, but parallel to the contact surface (the table) in pointing.

Modeling by Minimum Jerk

From our view that grasping is nothing more than pointing toward surfaces, it follows directly that any model that gives a good description of pointing toward surfaces will describe grasping equally well. We modified the minimum-jerk model for this purpose because it can be treated analytically. In describing pointing, the minimum-jerk model yields some systematic errors, especially in the velocity profile. We don't expect that a model for pointing will perform better on grasping. Therefore we are not surprised that these aspects of the predictions of the modified minimum-jerk model are also rather different from experimental observations on grasping, as Newell, Steenbergen, and Stelmach note. We already admitted this in the target article. We think, however, that the simplicity of the model is, for our purpose, more important than the range of effects it can handle. We will discuss these two issues—the quality of the predictions and the value of the model—in the next two sections. In this section we will discuss comments on the model itself.

An important aspect discussed by various commentators is that the minimum-jerk model is formulated in terms of the kinematics of the end-effector. This means that the model's predictions are independent of the choice of the parts of the body used to move the end-effector, and independent of the solution of any redundancy problem. Similar grasping behavior has indeed been observed using various end-effectors and solutions to redundancy problems, as argued for instance by Savelsbergh. However, when looking at details of behavior, motor control is more complex, as argued for instance by Morasso, Rosenbaum, and Stelmach. Shifting to a model based on the kinematical or dynamical properties of the limb (Morasso,

Rosenbaum) could help to explain behavior in such situations. However, such a model hides the general principles that determine the overall characteristics of grasping.

The minimum-jerk model differs from many other models in that the three orthogonal directions are treated independently. This is why we could start the derivation in the appendix regarding only one dimension. In Equations 1 and 2, the approach parameter therefore appears as a scalar. In the rest of the derivation, we assume that the approach parameter is a vector (misinterpreted by Rosenbaum) perpendicular to the surface of the disk. The angle φ is the direction of this vector, the parameter a_p its length. Rosenbaum argues that it is very unlikely that a_p is a controlled parameter. We do not claim that a_p is controlled. We consider the parameter a_p to be a tool for describing the behavior, not a control parameter.

Moreover, we explicitly do not claim that the brain minimizes jerk. We agree with Morasso that there is nothing special about the minimum-jerk model in generating smooth movements. We only claim that movements are smooth and tend to approach surfaces perpendicularly. In our view, the minimum-jerk model is a model of motor behavior, not one of motor control. The observed smoothness in movements is in our opinion the result of the interaction between various levels of motor control, from the cortical level to limb biomechanics. The minimum-jerk model describes the result of this interaction in a way that can be handled analytically. Several commentators (Morasso, Neilson, Rosenbaum) have formulated descriptions of mechanisms that could be responsible for the smoothness of movements, and others have been proposed elsewhere (e.g., Harris & Wolpert, 1998). Since our modeling effort is on the level of behavior, we will not discuss the pro's and con's of these proposed mechanisms. The minimum-jerk model is thus a simple description of the observed smoothness.

To incorporate our second claim—movements tend to approach surfaces perpendicularly—we had to modify the minimum-jerk model. Neilson questions why we use the final acceleration (and not velocity) to model these constraints imposed by the object. We have no real argument for this choice. In response to his question, we tested an alternative version of the model. We set the final acceleration to zero, and require that the final *velocity* be perpendicular to the object. The predictions of this modified model (shown in Figure 2) correspond qualitatively with our original model. Quantitatively, there are slight differences: the average slope of the relation between object size and maximum grip aperture is 0.88 instead of 0.81, and the maximum opening when grasping small objects occurs at 67% of the movement instead of at 60%.

Although this corresponds less well to the average of the experimental data, it is still within the range of experimental values found. This confirms that the model behavior is the result of our assumptions, not of the way we modeled them. We assume that any model (regardless of the level of description) which produces smooth trajectories that end more or less perpendicularly on the surface will give similar results. Rosenbaum shows that this is indeed true for his kinematics-based model.

We have formulated our predictions in terms of the development in time, because experimental results are generally presented in that form. As our model generates the complete movement kinematics, we can translate all results easily into spatial terminology, contrary to Stelmach's critique. In fact, our model is based on a spatial analysis of grasping, not on a temporal one.
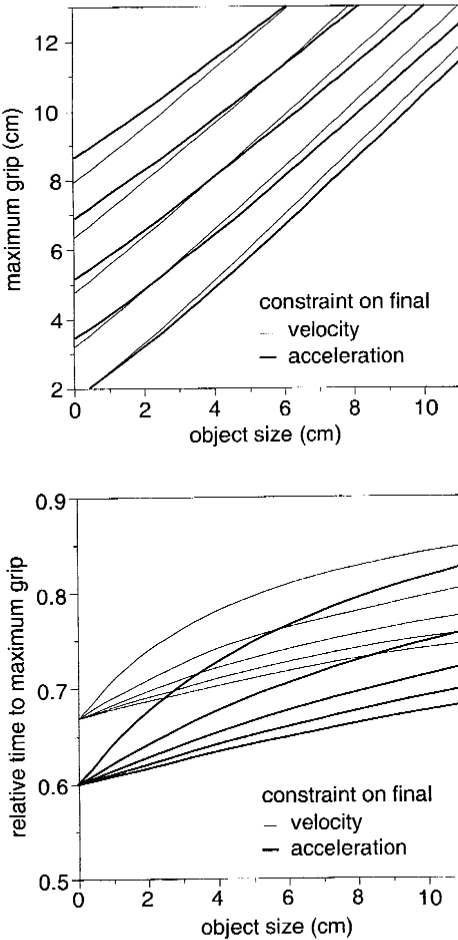
Figure 2 — Predictions for two types of constraints to incorporate the perpendicular approach in the minimum-jerk model. The format and thick lines are the same as in Figure 6 of the target article; the thin lines are the predictions based on a minimum-jerk model with a zero final deceleration and a final perpendicular velocity (scaled by MT) ranging from 0.4 to 2 m.

Prediction of Experimental Data

Weir remarks that the definition of the transport component we use in our target article differs from the definition on which most of the experimental values are based. In principle, this could indeed explain some of the differences between our model predictions and experimental data on the transport component. She concludes that the correspondence of the model predictions with the wealth of experimental evidence on the apparent independence of transport and grip is very attractive. Some other commentators, however, have their doubts.

Marteniuk and Stelmach argue that averaging the results of several independent studies is not without pitfalls. That is of course true. There is indeed a lot of variability in the experimental setup, number of subjects, object sizes, instructions, object shapes, marker placement, and so on. What may be surprising is that

these variations did not seem to have an effect on two relationships. In all experiments the maximum grip size and the time-to-peak aperture increase with object size. As our model predicts that experimental variations will have negligible effects on two of the regression coefficients of these relations, we simply averaged all experimentally found values for these two regression coefficients. As shown in Figure 7 of our target article, the range of parameters found for these relationships is rather limited, and seems to be distributed unimodally around an average value. Moreover, the 18 years of modern research on grasping have not revealed any way to influence these relationships systematically.

A problem with simply averaging regression coefficients is that the quality of the fits varies strongly. A solution would be to weight the contributions, as noted by Stelmach. This is in principle a good idea, but we do not have a simple formula for combining the number of disk sizes, repetitions, subjects, and information on other experimental parameters such as spatial and temporal resolution into one weight factor. Stelmach's suggestion to use r^2 as a weight factor is, in our opinion, not a very good choice. Experiments showing no effect (for instance of object size on the timing of maximum grip aperture) have a very low r^2 value, and would therefore not contribute to the average. Such a method would thus introduce an intolerable bias in the obtained average.

Marteniuk and Stelmach are in principle correct in their hesitance to simply average the data. Indeed, if we wanted to draw strong conclusions from a comparison with the data (e.g., "our model corresponds better to the data than model x"), we would definitely have to take more care of this issue. However, 18 years of research on grasping in the classical tradition has only resulted in one quantitative model (Hoff & Arbib, 1993). This model does not give predictions for the slopes and intercepts that our model predicts.

In discussing our choice of the minimum-jerk model, we already acknowledged that it could not explain the observed asymmetries in the transport velocity profile. Steenbergen argues further that this observed asymmetry depends on intrinsic object properties. We have no fundamental problems with a dependency of the transport component on intrinsic object properties. Remember that the independence we found only emerges from the model when finger and thumb have the same value for a_p. If the value for a_p is different for both digits, a slight asymmetry in the predicted velocity profile is predicted (see Figure 1). The observed dependency of the transport velocity profile on intrinsic object properties gives an additional illustration of our argument that the transport component cannot be part of a visuomotor channel based on extrinsic properties.

Stelmach brought in a last experiment that supposedly contradicts our model predictions. According to him, Kudoh et al. (1997) have shown that the development of the grip component is not independent of the reach distance. This is not our interpretation of their experiment. Our model predictions correspond perfectly to the behavior measured in the experiment by Kudoh et al. (1997): the average grip-parameters are independent of target distance. The only grip parameter that depended on target distance was the variability in the maximum aperture. This is consistent with our model: the speed of the digits increases with target distance, and faster movements are more variable (see Section 2.2 of the target article).

Savelsbergh discusses some recent studies on transport and grip coordination, from a category we deliberately did not discuss in our target article (Stelmach also mentions an unpublished experiment in this category). The common results in

these experiments is that transport and grip appear to be coordinated in the classical way, even though the grip formation is performed by another effector than the transport. This result appears at first to contradict our view on grasping, as claimed by Savelsbergh and Stelmach. As argued below, however, we think these results fit perfectly in our view, and we explain how.

A first experiment our model can explain is on mouth opening during eating (Castiello, 1997). It is not clear from that paper how the cheese was eaten, so we do not know exactly what the geometric constraints were. It seems plausible that when eating, we want to squeeze the cheese by our jaws. And to squeeze effectively, one has to approach the surface of the cheese orthogonally. To implement the model in a simple way, we assume $\varphi = 0$ (see Figure 3 of target article). For this situation, the movements of the jaws are purely in the x-direction and the transportation of the cheese is in the y-direction. Note that in the minimum-jerk model, perpendicular components of the movement are independent. The x-components of Equations 3 and 5 can thus describe the opening and closing movements of the upper and lower jaw when eating pieces of cheese of various sizes. For Castiello's (1997) experiment on eating, the model predictions are not only qualitatively correct but also quantitatively correct: the slope for the relationship between mouth opening and object size is 0.8.

The second experiment explained by our model is catching. When we regard the motion of the fingers relative to the ball, grasping the ball when it is stationary is very similar to catching it while it is moving. The only difference is that the component of digits' movements in the direction of the ball is passive in catching (the ball moves to the hand) but is active in grasping (the hand moves to the ball). The constraints on the positioning of the fingers are the same in both tasks.

We can implement the model in a similar way as for the eating experiment. A problem with catching experiments in the dark (as in the experiment by Van der Kamp et al., 1997) is that subjects have little information about the size and position of the ball and tend to use a default value. Formulated in terms of our model: the positions for contacting the ball are poorly defined, and subjects will start their movements toward default positions with a very high approach parameter. We expect therefore that the relationship between object size and maximal grip aperture will show a large intercept (large a_p) and a slope of less than 0.8 (effect of default position). As the lack of information is larger in the monocular condition, we expect a smaller slope and larger intercept. This is indeed the case: the slopes are 0.21 for monocular, 0.43 for binocular; the intercepts are 9.3 and 7.7 cm, respectively (Van der Kamp et al., 1997, Table 3). The experimental value for the intercept in the monocular condition (9.3 cm) is indeed very high, higher than reported in any other study.

The relationship between what happens after contact and the movement toward the object is clearly an important issue. Newell argues that our model should take this into account. We have demonstrated in our target article that with our model, two parameters could reflect the task constraints after contact. Both the approach parameter and the MT could depend on such constraints. In Sections 3.4 and 3.5.we gave several predictions on how the task after contact should influence behavior. Some of the predicted effects were found in the experiments, while no effect opposite to our predictions was found. We treated this as support for our model, though this support is rather weak, as Weir mentions. Perhaps subjects use

another strategy (outside the scope of our model) to deal with this aspect of the task. For instance, the digits' positions on the object presumably reflect the constraints imposed by the task after contact has been made. We hope that our model promotes experiments in which "the constraints arising from nested actions" (Newell) are varied. The explicit predictions our model makes can then be tested.

Value of Models

A model that can handle all possible experimental data is of no value. For instance, a neural network that learns by an error-backpropagation algorithm is extremely powerful in reproducing input/output relationships. However, as such a network can reproduce any relationship, it is of little use in understanding the relationship it has reproduced. A model should be restricted to be valuable. On the other hand, a model that can explain only one phenomenon is too restricted to be of great value. Several commentators question whether our model is restricted enough.

Both Savelsbergh and Newell argue in their final paragraphs that our model is nothing more than data fitting. Our model is based on a few principles, which are modeled using a minimum-jerk approach. This yields predictions of movement paths with only one variable: the approach parameter a_p. In discussing the wealth of experimental data, we concentrated on some general aspects, e.g., the relationship between maximum grip aperture and object size, which are independent of a_p. Thus we were testing a model that effectively had no parameter at all! We therefore do not think it is appropriate to summarize our work by "model fitting" (Newell) or "curve fitting" (Savelsbergh).

As our model is not based on an anatomical substrate, it cannot answer the question as to which muscles and joints are used to realize its solutions. Several commentators (Marteniuk, Rosenbaum, Newell) consider this limitation a major drawback. Steenbergen and Morasso make an even stronger claim. They argue that we can only learn from a model if it is related to the biological constraints. Our problem with this reasoning is that one does not know beforehand which aspects of biology limit a certain aspect of human performance. Many aspects of human motor performance remain invariant under changes in the effector system used (Merton, 1972). For instance, one's handwriting on a piece of paper and on a blackboard are quite similar, despite the different set of muscles and joints used in these tasks. The main determinants in these movements are thus not in the muscles and joints but elsewhere in the (neuro)biology.

Our idea about modeling motor behavior is that it should help us understand how this behavior emerges. To achieve this, the first requisite for a model is that it be much simpler than the motor control apparatus it describes. The easiest way to achieve this is to omit elements that do not determine the behavior. By making the model as simple as possible, one obviously loses some of the versatility of the human motor system. To the extent that biological constraints are limiting, they should emerge from this approach.

Marteniuk asks himself in his first paragraph what could disprove our model. That question is easy to answer: it is already disproved, for instance by the asymmetry in the velocity profile of the transport component. However, we do not claim to propose a perfect model but to propose a view. We deliberately chose the simplest model we could think of to make some quantitative predictions. In a recent grant proposal, we suggested several experiments that could disprove our view.

Savelsbergh proposes a nice experiment in this spirit. However, experiments are not very well suited for disproving views. The history of experiments on grasping shows it is always possible to introduce additional assumptions to keep a view alive (e.g., by accepting interactions between independent channels). More generally, we do not think that one can disprove views. One can only try to convince others that a view is not fruitful, or that it is not elegant.

References

Brenner, E., & Smeets, J.B.J. (1995). Moving one's finger to a visually specified position: Target orientation influences the finger's path. *Experimental Brain Research*, **105**, 318-320.

Brenner, E., & Smeets, J.B.J. (1996). Size illusion influences how we lift but not how we grasp an object. *Experimental Brain Research*, **111**, 473-476.

Carnahan, H., Goodale, M.A., & Marteniuk, R.G. (1993). Grasping versus pointing and the differential use of visual feedback. *Human Movement Science*, **12**, 219-234.

Castiello, U. (1997). Arm and mouth coordination during the eating action in humans: A kinematic analysis. *Experimental Brain Research*, **115**, 552-556.

Cole, K.J., Gracco, V.L., & Abbs, J.H. (1984). Autogenic and nonautogenic sensorimotor actions in the control of multiarticulate hand movements. *Experimental Brain Research*, **56**, 582-585.

Cole, K.J., & Abbs, J.H. (1986). Coordination of three-joint digit movements for rapid finger-thumb grasp. *Journal of Neurophysiology*, **55**, 1407-1423.

Harris, C.M., & Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. *Nature*, **394**, 780-784.

Hoff, B., & Arbib, M.A. (1993). Models of trajectory formation and temporal interaction of reach and grasp. *Journal of Motor Behavior*, **25**, 175-192.

Jeannerod, M., Arbib, M.A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: The cortical mechanisms of visuomotor transformation. *Trends in Neuroscience*, **18**, 314-320.

Klein Breteler, M.D., Meulenbroek, R.G.J., & Gielen, S.C.A.M. (1998). Geometric features of workspace and joint-space paths of 3D reaching movements. *Acta Psychologica*, **100**, 37-53.

Kudoh, N., Hattori, M., Numata, N., & Maruyama, K. (1997). An analysis of spatiotemporal variability during prehension movements: Effects of object size and distance. *Experimental Brain Research*, **117**, 457-464.

Lemon, R.N., Johansson, R.S., & Westling, G. (1995). Corticospinal control during reach, grasp, and precision lift in man. *Journal of Neuroscience*, **15**, 6145-6156.

Merton, P.A. (1972). How we control the contraction of our muscles. *Scientific American*, **226**(5), 30-37

Paulignan, Y., & Jeannerod, M. (1996). Prehension movements: The visuomotor channels hypothesis revisited. In A.M. Wing, P. Haggard, & R. Flanagan (Eds.), *Hand and brain: Neurophysiology and psychology of hand movements* (pp. 265-282). Orlando: Academic Press.

Steenbergen, B., Hulstijn, W., Lemmens, I.H.L., & Meulenbroek, R.G.J. (1998). The timing of prehensile movements in subjects with cerebral palsy. *Developmental Medicine and Child Neurology*, **40**, 108-114.

Van der Kamp, J., Savelsbergh, G., & Smeets, J. (1997). Multiple information sources guiding the timing of interceptive actions. *Human Movement Science*, **16**, 787-822.