Universität Bielefeld
Faculty of Psychology and Sports Science
Neurocognition and Action - Biomechanics
Dr. Christoph S c h ü t z

Universität Bielefeld # Postfach 10 01 31 # 33501 Bielefeld

Dr. Melvyn A. Goodale
Editorial board
Experimental Brain Research

Phone: (0521) 106 - 5128
Fax: (0521) 106 - 6432
@Mail: christoph.schuetz@uni-bielefeld.de
www.neurocognition.de / www.cit-ec.uni-bielefeld.de

Bielefeld, 18 December 2015

Manuscript Revision

Dear Dr. Goodale,

Please find enclosed the revision of the manuscript EXBR-D-15-00556 entitled "Cognitive costs of motor planning do not differ between pointing and grasping in a sequential task" by Thomas Schack, Matthias Weigelt, and me.

We wish to thank you for handling the manuscript, and we wish to thank both reviewers for their thoughtful and helpful comments on the original submission. In the revision, we have attended to every comment made by the reviewers. Below we reproduce the original comments by the reviewers followed by our responses, highlighted in orange. In our responses, we indicate the pages in the manuscript where changes were applied with regard to the reviewers´ comments. Changes to the manuscript are highlighted in coloured text.

We hope that this revision is now suitable for publication in Experimental Brain Research, and look forward to hearing from you about the disposition of the revision. Thank you very much for your kind attention.

This letter and the revision have been read, contributed to, and approved by all of my co-authors.

Yours Sincerely,

Dr. Christoph Schütz

Reviewer #1

Overall, this is a very-well written and relatively straightforward paper reporting an interesting experiment on an interesting topic. This reviewer thinks the manuscript is suitable for publication in EBR.

We thank the reviewer for the positive evaluation of the manuscript.

This reviewer has mostly minor comments. My only 'major' issue concerns the limitations of the current analysis.

Specific points as follows

The authors use motor hysteresis as a proxy index of the cognitive cost of movement planning. This reader thinks the abstract would be improved if the authors included a sentence conveying the direction of that relationship as they report it in the introduction: more hysteresis = less cognitive cost.

The reviewer raises a valid point. We have added this information to the abstract (see page 1) .

How can movement planning costs be the same between pointing and grasping (as the authors assert) despite a greater range of motion for grasping movements? It suggests, to me at least, that operationalizing movement-planning cost in terms of only hysteresis is problematic. Perhaps the authors mean to say that they suppose that hysteresis affects grasping and pointing movements equally. At any rate, a bit more detail would be helpful here.

We think that movement planning costs depend on the dimensionality of the solution space. As long as both tasks can be solved with the same number of independent degrees of freedom, the planning costs should be similar. Both tasks in the current study could be solved with the same number of independent degrees of freedom, as stated in the discussion (see page 9). The ranges of motion used for each degree of freedom do not affect the cognitive costs of motor planning.

It would be less confusing for the reader if the authors avoided the term "effect size" of observed hysteresis. "Effect size" has a fairly specific statistical meaning. In the discussion, the authors mention "the smaller hysteresis effect..." which seems to this reader a more appropriate phraseology. Or perhaps, "the size of the hysteresis effect"?

We thank the reviewer for this remark. We agree that "effect size" has a specific statistical meaning and, thus, could lead to confusion, especially as the statistical analyses used in the manuscript are complex enough. The term "hysteresis effect size" has therefore been replaced by "the size of the hysteresis effect" or "the hysteresis effect" throughout the manuscript. We agree with the reviewer that this change has improved the overall clarity of the manuscript.

Can the authors elaborate a bit more on why reuse increases the mechanical cost of a movement? Does this statement apply to an iterative action performed on the same object? Or is this statement applying to sequential actions performed on different objects? I assume the latter, but the authors could specify one way or the other to resolve the ambiguity.

The statement applies to the upcoming movement in a sequential task. More reuse results in a persistence on the former, less optimal motor plan, which increases the mechanical cost of motor execution. We have added this information to the introduction (see page 2).

Can the authors specify what they mean when they write "...an increase of the mechanical cost of the task for ten sequences significantly reduced hysteresis effect size." Given what the authors report earlier in the manuscript, I was under the impression that increasing hysteresis (a product of "reuse") is associated with increases in mechanical costs.

The reviewer is certainly correct when stating that an increase in plan reuse is associated with an increase in mechanical cost (see Fig. 1a). This effect, however, is only theoretical, as plan reuse cannot be manipulated experimentally. Instead, an optimal fraction of reuse is automatically set based on the mechanical and cognitive cost of the task (see Fig. 1a).

However, as stated in the introduction, the coupling of the cost factors and the fraction of reuse is bidirectional (see page 2). If the mechanical cost of the task is increased, a new optimal fraction of reuse is set automatically, which once again minimises the total cost of the movement. Importantly, the new fraction of reuse is lower. This results in a slight increase of the cognitive cost, which is rewarded by a larger decrease of the mechanical cost (resulting in a new stable state which would look like a mirror image of Fig. 1b). Or, in more practical terms, people use a more optimal posture, which requires more cognitive cost, but is rewarded by a facilitated motor execution.

We have extended the paragraph on the bidirectional effect to clarify this issue (see page 2).

Why did the authors choose two different target stimuli for the grasping and pointing tasks? Also, it isn't clear to this reader where exactly the target sphere is located. At the moment, this reader is guessing that the recess for the sphere was positioned at the side of the knob that faced the participant. At any rate, it would be helpful if the authors updated Fig. 2b. to point out the sphere.

We thank the reviewer for this comment. We assumed the location of the pointing target to be evident from the figure. As this does not seem to be the case, we have updated Fig. 2 with pointers to the target sphere and the drawer handle (see page 4). There are no two target stimuli for grasping and pointing. The target sphere is integrated into the drawer handle. We have further rewritten the methods section to better convey this idea (see page 4).

The authors hypothesize that when participants perform the grasping task, they become sensitized to the tight constraints on hand orientation and increase the constraint's rank in the hierarchy. This reader cannot understand why this alone would predict an effect on the subsequent pointing behavior as the authors claim in the sentence that immediately follows. Why would a constraint's rank for one task transfer to another? The authors should bolster the connection between their hypothesis and this prediction by filling in the missing detail. Reuse?

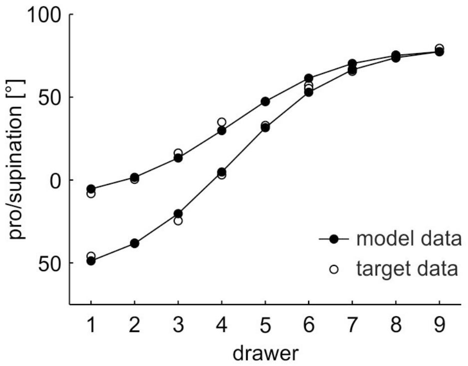
We apologize for not clarifying this enough. Indeed we think that participants would reuse the former constraint hierarchy as long as this does not increase the mechanical costs of the subsequent task significantly. We have rewritten the mentioned paragraph in order to convey this concept more clearly (see page 3/4).

The authors analyze the drawer x task interaction presupposing that a range effect and only a range effect would best characterize the observed distinction in the pro/supination curves between tasks. I don't take issue with the hypothesis that drives this prediction, but I do take issue with the absence of any tests on the slopes and offsets of the sigmoidal functions - the authors are upfront about the assumptions under which the test of their hypothesis would be valid, and it seems to me that should test these assumptions. In Fig. 4, for example, it doesn't look like the slopes are similar. Furthermore, the authors post-hoc per-drawer analysis does not fit the notion that a restriction of range and only a restriction of range is operating here.

The reviewer raises a valid point. We would kindly ask the reviewer to first read the reply to his/her next question ("This leads me to a more general..."), where we provide an in-depth analysis on the use of a sigmoid function for modelling, the parameters involved, and the problems of statistical analyses conducted on these parameters. Regarding the specific issue raised here, the reviewer is correct when stating that the slopes of both functions in Fig. 4 are dissimilar. This is due to a miscommunication on our side. When stating that 'the interaction pattern followed a trend resulting from two sigmoid functions with the same slope', we referred to the specific model parameter:

y = **y_offset**+**y_range***0.5*(tanh(**slope***(x-**x_offset**)))

Indeed, the measured values for the pointing and grasping task can be fitted quite nicely by two sigmoid functions which use the same **slope** and **x_offset** parameters, but a different **y_range** parameter:



The differences in the overall gradient can result solely from the change in the **range** parameter. Of course, the **y_offset** parameter in this case was also adjusted, to achieve a similar upper boundary value (which reflects the pattern of results found in the real data, i.e., an extension of the range by a shift of the lower boundary). For the contrast analysis, the **y_offset** parameter is actually irrelevant, as the mean of the coefficients is normalized to zero, and any difference in the **y-offset** is eliminated from the coefficients in this step.

To address the issue raised by the reviewer, we never wanted to claim that there is no difference in the slope or offset of the two functions, but only that a change in range would be sufficient to explain the interaction (which it is, as demonstrated by the contrast analysis). To clarify this, we have added a footnote that slope refers to the model parameter only and further added a statement that the contrast indicates that a change of the range can be sufficient to explain the interaction (see page 7/8).

Regarding tests on the slopes and offset, these are not possible due to the problem of parameter averaging mentioned in the subsequent reply.

This leads me to a more general recommendation. Why not submit the participant-specific parameters that fall out of fitting a sigmoidal function to each participant's data set to a rmANOVA? Surely there are desirable members of this family of functions that have parameters that can be easily mapped onto the features of the response curves (e.g., range, slope, and horizontal shifts) the authors are interested in studying. An analysis of this nature would more fully describe the relationships between the manipulated variables and the resultant postures, and put the authors in a better position to qualify their current interpretation.

The reviewer raises a fascinating point. Indeed, if the data could be represented by the parameters of a sigmoid function, a much more clear-cut analysis of the effects of each factor on the range, slope, and horizontal offset would be possible. We have therefore conducted the suggested analysis. While we were able to fit the current data set with a sigmoid function, the resulting parameters cannot be analysed by an rmANOVA. We would like to demonstrate this in more detail:

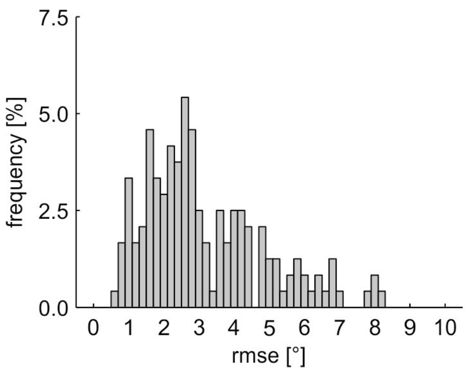
The sigmoid function used for the modelling was based on a hyperbolic tangent function. To render this function suitable for representing the data, four adjustable parameters were required:

x_offset = left-right movement of the model function
y_offset = up-down movement of the model function
y_range = range of the model function
slope = slope of the model function

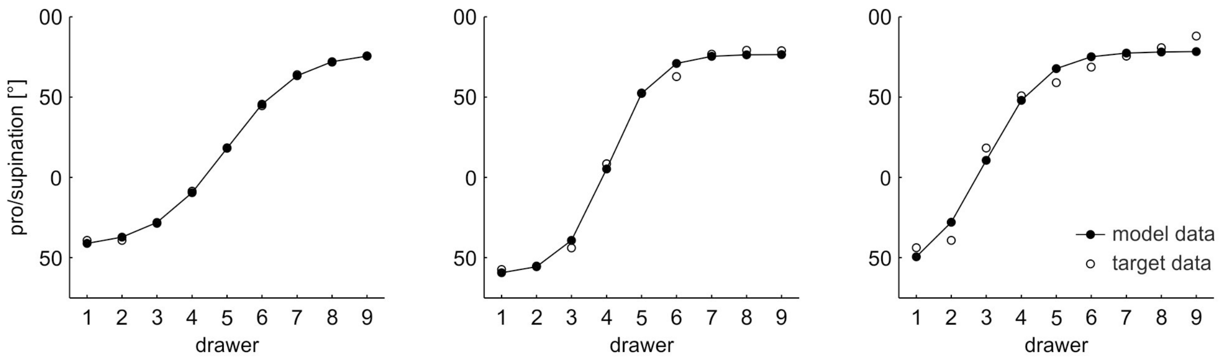
The model output was calculated as follows: y = **y_offset**+**y_range***0.5*(tanh(**slope***(x-**x_offset**)))

The drawer heights (x = 1 - 9) were used as input. As the tanh function has a range from -1 to +1, a multiplier of 0.5 was necessary to make **y_range** reflect the real range of motion. Further, while the **slope** variable indeed represents the slope of the sigmoid function, the total derivative also depends on the **y_range**, which stretches the output function. This effect, however, applies to any function from the family of sigmoid functions and, thus, cannot be avoided.

To reduce the variance of the fitted data, we used the average of the four repetitions conducted for each factor combination. Please note that the factors no longer include 'height', as the joint angles for the nine target heights were used as the target values for the model. An individual fit was calculated for each participant, sequence, and movement task. To evaluate the quality of the fit, the root mean squared error (rmse) between the target values and the model values was calculated. Distribution of the rmse-values is depicted in the following histogram:

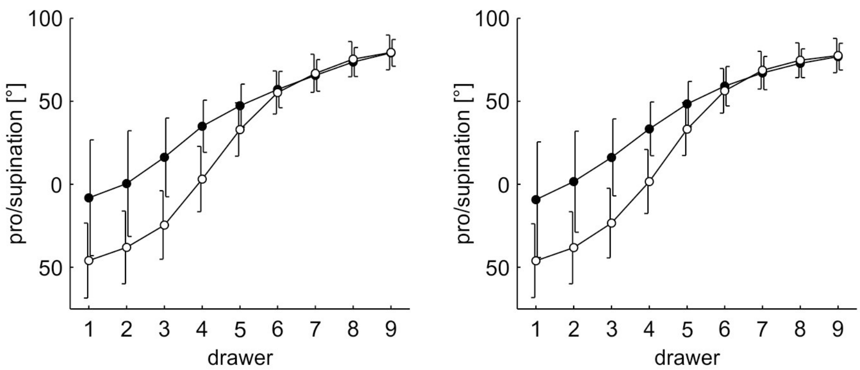


Please note that the maximum deviation is 8.1° and that 66% of the sequences have an rmse-value of less than 3.6°. Taking the range of motion of over 100° into account, we considered the model fit acceptable. To give the reviewer an impression of the individual model performance, we had MATLAB randomly select one trial from the lower, central, and upper rmse-tertile (shown from left to right) and depict them together with the model fit:

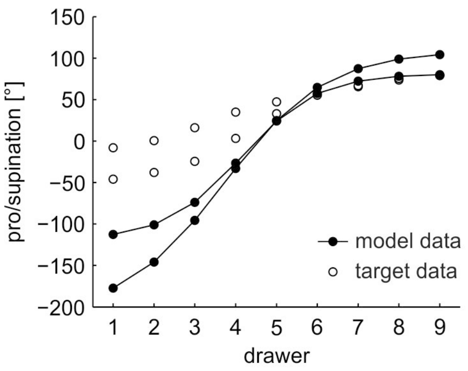


The graph on the right demonstrates how the model fit can result in an overestimation of the **y_range**. In our (subjective) opinion, the fitted model has not captured the lower boundary value correctly. The derivative should be close to zero between drawers 1 and 2. However, the value at drawer 2 might as well represent an outlier. For the analysis, we had to rely on the objective estimation of the parameters and, thus, used the output from the least-squares optimisation algorithm implemented in MATLAB.

We tested for an effect of the factors sequence and movement task on the rmse-values and found a significant main effect for movement task, $F(1,38) = 5.921$, $p = .020$. Grasping movements on average had a larger rmse-value, which can be attributed to the larger range of motion. To verify that the averaged model output reflected the averaged target values, we recreated the 'data split by movement task plot' (Fig. 4 in the manuscript) once with the measured (left) and once with the modelled (right) data:



Similar figures were created for group and sequence. The modelled data in all cases looked almost identical to the real data, with only minor differences in means and variances. However, these plots were created by averaging the individual model outputs for each participant and factor combination. To apply an rmANOVA to the <u>parameter data</u>, the sigmoid function created from the <u>averaged parameters</u> of the individual models would need to reflect the real data as well. This, however, is not the case. The following plot depicts the target and model data created by an averaging of the parameters. The target data used is the one from the previous plot:



The model data deviates completely from the target data. The sigmoid functions created by an averaging of the parameters do not represent the real data. As ANONAs are based on column means, an rmANOVA applied to the parameter values would output results which did not reflect the real differences in posture.

We have therefore not implemented this analysis in the revised version of the manuscript. The currently used contrast analysis conducted on the means over all participants to us still seems to be the most valid approach which takes the sigmoid trend of the data into account. We hope that the reviewer agrees with this decision.

A minor point, but the authors should change their wording from "confirm" to "support" when relating their findings back to their hypothesis.

We have replaced "confirm" by "support" on all occurrences relating to the hypothesis.

Reviewer #2

In this work, the authors present an experiment aimed at establishing whether cognitive costs differ between two types of movements, namely grasping and reaching.

The experimental setup is a modification of one previously used by the authors, composed by a set of 9 drawers aligned in vertical, with the participant standing in front of them and either grasping the drawer's knob of reaching to a ball in the center of the knob. Order of movements has been randomized across participants. Movement data are recorded and kinematics parameters are normalized based on the physical ones. The dependent variable to explore the authors hypothesis is the global pro/supination angle calculated.

When taking into account results, together with the type of movement the authors take into account the drawer position (lowest or highest), the sequence of movements (ascending or descending exploration of the vertical set of drawers),the order of movements (i.e. first grasping). According to the authors results, cognitive costs do not differ for the two types of movements, so previously found effects should be attributed to the spatial configuration adopted in the task.

In general, the methodology to investigate the question pushed forward by the authors is well designed. Nonetheless, there are some concerns regarding the reliability of the results, as the number of variables in the game seriously undermine the statistical validity. Further the introduction does not convey the aim of the study due to its complexity. Consequently, some modifications are required to strengthen the manuscript - please see detailed comments below.

We thank the reviewer for the positive evaluation of the methodology. The mentioned concerns are addressed in the "detailed comments" section.

Finally, it would be beneficial for the authors to include a more widen view of how they results expand knowledge on movements, behind the specific experiment and maybe referring a bit more on the practical implications of such a result - i.e. if the movements have the same cognitive costs, how doe this apply to movement research in general? Is there any impact related to neural basis of movement control? are there implications maybe in prosthetic research as well?

The reviewer raises a valid point. However, we would like to refrain from drawing extensive conclusions at this juncture. While the cognitive cost of pointing and grasping does not differ in the current task, this might be a result of the (vertical) target configuration. As mentioned in the discussion, the cognitive cost of motor planning could be linked to the dimensionality of the joint angle solution space. The dimensionality might be the same for both pointing and grasping in the current target configuration, as two independent degrees of freedom could be sufficient for both tasks. We are planning on a follow-up paper in which the dimensionality of the solution space is analysed in detail and the wider significance of the current results is assessed. Both analyses were originally presented in one manuscript, which was rejected due to its excessive length. We were therefore forced to split it for publication. The implications of our results will be discussed in the second paper.

Detailed comments:

1) Introduction

The introduction is long and it is difficult to focus on the aim of the study, which is clearly explained in the beginning of the discussion and in the abstract. Many theories, concepts and related experiments are exposed in the introduction, and the reader finds it difficult to follow the path towards the main point of the study. The background for the study would benefit from a reorganization of relevant information, with more focus on the main aim (cognitive costs in reaching and grasping) and less focus on possible secondary effects. Priority shall be given to the theories more related to the experiment itself.

We thank the reviewer for this remark. We have shortened the introduction considerably and now focus only on motor hysteresis as a proxy for the cognitive cost of a task and studies required to derive the hypotheses at the end of the introduction.

2) Methods

In the methods section many sentences are exactly the same or too similar as in the authors previous paper in Experimental Brain Research 2013 Influence of mechanical load on sequential effects. See for instance line 23 to 41 at page 5, line 26 to 44 at page 6 and figure 2 which is only slight modified form the previous work. Maybe it is better to refer to the previous work instead of reporting the same information, specifying only the setup modifications.

The reviewer raises a valid point. We have therefore condensed the "preparation" section (formerly line 23 to 41 at page 5) and refer to the EBR paper for details (see page 4/5). Table 1 has been removed. In the "kinematic analysis" section (formerly line 26 to 44 at page 6), the first paragraph has been rewritten and condensed (see page 5). The second paragraph, however, has not been modified, as there are important changes with respect to the previous paper (the hand vector has been replaced by a wrist vector, which is better suited for the analysis of a pointing and grasping task). The same argument applies to Fig. 2, which has also been extended by additional labels for the pointing target and the grasping handle, as requested by reviewer #1.

Secondly, it is somehow hard to follow the description of the trials - for instance stating that there are eight real test sequences in the beginning of the description and then specifying nine trials an only later on what sequences are (ascending and descending movements) is a bit confusing. Similar confusion arises from page 6 line 17 on. It would be better to describe immediately what are sequences, trials etc, before the procedure.

We thank the reviewer for this remark. We have reordered this section to put a definition of task, sequence, and trial at the beginning. We have further rechecked the phrasing to be consistent with these definitions (see page 5). We agree with the reviewer that the restructuring has improved the readability of the manuscript.

On the same line, at page 6 the authors state they analyze 144 angles - why not simply stating that the 72 trials of the test phase have been analyzed?

We have rewritten this paragraph to be less confusing (see page 6). Please note that indeed 144 trials from the test phase were analysed, corresponding to 72 trials from the pointing and grasping task, respectively.

3) Results

The main concern about results is the related to the small effect sizes reported for the effects. The impression is that this is related to the complexity of the experimental design, including a number of levels for the factor drawer plus other 3 factors within and between. The factors included in the analysis do not appear however totally related to the main experimental question - so the question is whether it would be more effective to run an analysis centered on the experimental question and secondarily run an analysis exploring the factors that could also affect the results. I appreciate the authors consider all the variables in the beginning, but the statistical power looses a lot this way.

The reviewer is certainly correct that the effect sizes for a majority of the factors are small. However, please note that η^2 values (not partial η^2 values) are reported. The small effect sizes most likely result from the overwhelming fraction of variance that is taken up by the factor 'drawer' (73.33%). As 'drawer' is a required factor in all the interaction analyses in the manuscript, running analyses centred on a reduced number of factors would yield no different results from the full-factorial rmANOVA. Reporting partial η^2 values, on the other hand, would.

The reviewer asked us to first run an analysis focused on the main question. Regarding the main experimental question - the difference in the size of the hysteresis effect between pointing and grasping - a reduction of the factors included in the rmANOVA does not improve the analysis. To show an effect of movement task on the size of the hysteresis effect, we need at least the factors 'movement task' and 'sequence'. If we conduct a focused 2 ('movement task') × 2 ('sequence') rmANOVA on the data, we get the same *p*-value for the interaction as in the full rmANOVA, $F(1,39) = 0.893$, $p = .350$, $\eta^2 < .001$. As we already know from previous studies that a hysteresis effect is mostly restricted to the central drawers, removing the factor 'drawer' and averaging across drawers (including the outer ones without an effect) could actually conceal a positive effect. Indeed, after adding 'drawer' as a factor, the *p*-value for the interaction is closer to significance, $F(8,312) = 1.861$, $p = .124$, $\eta^2 < .001$.

For the subsequent analyses, effect sizes do not improve by a reduction of the factors: For the interaction of 'sequence' × 'drawer', a focused 2 × 9 rmANOVA would give us $F(8,312) = 15.743$, $p < .001$, $\eta^2 = .003$ (compared to .002 for the full-factorial rmANOVA). For the interaction of 'movement task' × 'drawer', a focused 2 × 9 rmANOVA would give us $F(8,312) = 83.807$, $p < .001$, $\eta^2 = .040$ (compared to .040 for the full-factorial rmANOVA). A similarly negligible difference in effect size between the full-factorial rmANOVA and the focused rmANOVA is found for the three-way interaction of 'movement task' × 'drawer' × 'group'. Thus, reducing the number of factors in the rmANOVA does not affect the effect sizes as long as the factor 'drawer' is part of the analysis.

As mentioned beforehand, the use of partial η^2 values would change the outcome considerably. Let's repeat the paragraph on all significant effects with η^2_p instead of η^2:

The main effects of 'movement task', $F(1,38) = 91.353$, $p < .001$, $\eta^2_p = .706$, 'sequence', $F(1,38) = 88.508$, $p < .001$, $\eta^2_p = .700$, and 'drawer', $F(8,304) = 450.320$, $p < .001$, $\eta^2_p = .922$, were significant. There were significant two-way interactions of 'movement task' × 'drawer' (hybrid), $F(8,304) = 89.390$, $p < .001$, $\eta^2_p = .702$, and of 'sequence' × 'drawer' (pure ordinal), $F(8,304) = 15.668$, $p < .001$, $\eta^2_p = .292$, as well as a significant three-way interaction of 'movement task' × 'drawer' × 'group' (hybrid), $F(8,304) = 3.598$, $p = .031$, $\eta^2_p = .086$.

However, we prefer to report η^2 values, which, in contrast to partial η^2, reflect the percentage of total variance accounted for by each factor in the model. Especially with regard to the subsequent contrast analyses, in which we report the fraction of each interaction's variance that is accounted for by the contrast pattern, η^2 seems a more consistent choice for the reported effect sizes. We hope the reviewer can agree with this argument.

Secondly follow up analysis in this section (i.e. analysis of the interactions) are not detailed, i.e. no statistical values are reported for instance at the end of page 7.

The reviewer raises a good point. We have not reported the statistical values for the 2 × 9 rmANOVAs, as these only serve as a reference to calculate which fraction of the interaction is accounted for by the contrast pattern. As demonstrated in the previous reply, the results of a 2 × 9 rmANOVA do not differ significantly from those of the full-factorial analysis. However, as this might not be clear to the reader, we have added the ANOVA result for each interaction to avoid misunderstandings (see page 7/8).

Finally at page 9 the authors report their predictions, but they would be a better fit for the beginning of this section.

We have moved the predictions regarding the main result to the beginning of the section (see page 6).

4) Discussion

From page 10 line 10 secondary effects are explained, while the main discussion of the results is at page 11. The authors could consider moving the discussion of secondary effects, and go directly into the main point at the beginning of the discussion.

We thank the reviewer for this suggestion. The discussion of the main effect has been moved to the beginning of the discussion section (see page 9).

Secondly, the authors could consider that while they deeply discuss implications of their study in terms of cognitive costs as measured by pro/supination angle, there are in fact many other measures that could lead different results. Especially considering programming costs could arise before movement performance, i.e. in the planning rather than online phase, measures such as reaction times and movement velocity could be sensitive enough to highlight differences. Instead of assuming that no differences exist between the movements, the authors could specify that these differences are not visible in terms of pro/supination angle measurements.

The reviewer is certainly correct that other proxy measures of motor planning could yield different results. While we are unsure how movement velocity could serve as a measure of cognitive cost, programming costs arising before movement performance could be quantified by a measurement of movement initiation time. In the current study, however, we cannot calculate movement initiation time. Participants in the ordered sequences of trials can plan for all upcoming trials either in advance or between trials and, thus, the specific moment of planning onset cannot be defined. We have, however, added a comment on the restrictions of the current analysis to the discussion section (see page 10).