Article:

Mylon, P., Lewis, R., Carré, M.J. et al. (1 more author) (2016) An evaluation of dexterity and cutaneous sensibility tests for use with medical gloves. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 230 (16). pp. 2896-2912. ISSN 0954-4062

https://doi.org/10.1177/0954406215604005

An Evaluation of Dexterity and Cutaneous Sensibility Tests for Use with Medical Gloves

Peter Mylon[a], Roger Lewis[a], Matt J Carré[a], Nicolas Martin[b]

[a]Department of Mechanical Engineering, University of Sheffield, Sheffield, UK

[b]School of Clinical Dentistry, University of Sheffield, Sheffield, UK

Dr Peter Mylon – Corresponding Author
Eurolens Research
The University of Manchester
Carys Bannister Building, Room 1.017
Dover Street
Manchester
M13 9PL
(0161) 2755574
peter.mylon@manchester.ac.uk

Text word count: 7,441

References word count: 374

Running head: Glove Test Evaluation

Précis

An evaluation of two dexterity tests – the Purdue Pegboard Test and the Crawford Small Parts Dexterity Test – and a cutaneous sensibility test – the Semmes-Weinstein Monofilaments – for the assessment of the effect of medical gloves on manual performance. Only the Crawford 'Screws' Test showed significant differences in performance between glove types.

An Evaluation of Dexterity and Cutaneous Sensibility Tests for Use with Medical Gloves

Peter Mylon[1], Matt J Carré[1], Roger Lewis[1], Nicolas Martin[2]

[1]Department of Mechanical Engineering, The University of Sheffield

[2]School of Clinical Dentistry, The University of Sheffield

Objective: The ability of selected dexterity and cutaneous sensibility tests to measure the effect of medical glove properties (material, fit, and number of layers) on manual performance was analyzed.

Background: Manual performance testing of gloves to-date has focused on thicker gloves where the effects are more obvious. However, clinicians have reported dissatisfaction with some medical gloves and a perceived detriment to performance of new materials compared to latex.

Method: Three tests (Purdue Pegboard Test, Crawford Small Parts Dexterity Test and Semmes-Weinstein Monofilaments) were performed by 18 subjects in five hand conditions (ungloved; best-fitting, looser-fitting and a double layer of latex examination gloves; best-fitting vinyl gloves). Tests were performed in the ungloved condition first, and the order of the gloved tests was randomized. Learning behavior was also measured.

Results: The Purdue test showed a significant effect of hand condition, but no differences between latex and vinyl. No significant effect of hand condition was found in the Crawford 'Pins and Collars' test, but the 'Screws' test showed promising discrimination between glove types. The Monofilaments test showed a significant effect of hand condition on cutaneous sensibility, particularly a reduction when 'double-gloving', but no significant differences between glove types.

Conclusion: Existing tests show some ability to measure the effect of gloves and their properties on manual performance, but are not comprehensive and require further validation.

Application: In order to fully describe the effects of medical gloves on manual performance, further tests should be designed with greater resolution, and that better replicate clinical manual tasks.

Keywords: manual performance, motor skill, Purdue, Crawford, monofilaments

INTRODUCTION

Many dexterity tests have been designed, mostly with the aim of assessing the motor skills and hand-eye co-ordination of potential employees or for aiding in the rehabilitation of patients with brain or motor injuries. The earliest attempts to identify glove effects on dexterity (Griffin, 1944) used a cribbage board with cold-weather military gloves to determine the extent to which the gloves were impeding the performance of manual tasks, and many studies since have focused on similarly thick gloves. In these cases, the reduction in performance from the norm can be fairly substantial. Medical gloves, which are essentially a thin, flexible membrane, tend to have a much smaller, and hence less measurable, effect on dexterity. However, perceived differences in performance and dissatisfaction with certain types of glove amongst clinicians has been documented (Mylon, Lewis, Carré, Martin, & Brown, 2013).

The same can be said of cutaneous sensibility (the ability to sense external stimuli through the skin). The main use of apparatus such as the Semmes-Weinstein Monofilaments has been to measure loss of sensitivity due to nerve damage and assess the rehabilitation of patients after a stroke, for example. The loss of sensitivity caused by a thin rubber membrane is likely to be significantly less. In order to select the most appropriate tests for medical glove evaluation, it is therefore necessary to validate the tests in terms of ability to identify performance differences between gloves, repeatability and relevance to medical practice.

Previous work, consisting of a review of test methods relevant to glove design (Mylon, Carré, Lewis, & Martin, 2011) and interviews with practitioners in which the most manually-demanding tasks were identified (Mylon, et al., 2013), was combined to select the most appropriate tests for more detailed evaluation. Tests were selected using the Weighted Scoring Method, taking into account factors such as: cost and availability; ease and duration of procedure; proven ability to discriminate between gloves; and application to medical practice. Two standard dexterity tests – the Purdue Pegboard Test and the Crawford Small Parts Dexterity Test (CSPDT) – and one standard cutaneous sensibility test – the Semmes-Weinstein Monofilaments Test – were chosen for further validation with medical gloves. The two dexterity tests assess different manual skills – the Purdue Pegboard uses fine finger dexterity, while the Crawford test uses fine tool dexterity. Both skills are required in medical practice.

A recent study by Johnson et al. (2013) evaluated medical gloves using the same three tests. They concluded that, while there was a clear reduction in performance when wearing gloves compared to bare hands, performance did not vary significantly between gloves. However, since the focus of the study was on glove comparison, rather than test evaluation, the relevance of the performance measures to clinical practice and the ability of the tests to detect clinically relevant differences in performance were not addressed. Furthermore, no attempt was made to isolate the effects of variables such as glove fit and material, to assess the effect of multiple layers on manual performance, or to explain the results in terms of glove design.

This study was designed, as far as possible, to test glove variables independently, in order to draw rigorous and useful conclusions for glove design and selection. However, the primary purpose of the study was to validate the three tests methods as tools for glove evaluation. Therefore, based on the results of the study, the merits and shortcomings of each of the test methods in this regard are also discussed.

METHODS

All of the test protocols for the project, along with participant information sheets and consent forms, were submitted to the University of Sheffield Research Ethics Committee and received approval. None of the tests performed require any previous experience or specialized skills – they can be performed by anyone with a degree of manual dexterity.

Subjects

18 volunteers took part in the tests. They were all students at the University of Sheffield between 21 and 30 years of age. 16 of the subjects were male and two were female. They were required to be generally healthy and have no known sensorimotor deficiencies.

Gloves

The gloves used were POLYCOHealthcare (BM Polyco Ltd, Enfield, UK) ambidextrous examination gloves. Two types were used in this study: Finex® PF (powder-free) latex gloves, which are chlorinated on the outside surface to reduce allergens and coated with a polymer on the inside to improve donning; and Finity PF (powder-free) vinyl gloves. Each of the gloves had five available sizes: Extra-Small (XS), Small (S), Medium (M), Large (L) and Extra-Large (XL).

Variables

The independent variable, or 'within-subjects factor', in all the tests was hand condition, consisting of five levels: 'No Gloves', 'Best-Fit Latex', 'Best-Fit Vinyl', 'Double Best-Fit Latex' and 'Larger Latex'. These conditions allowed for analysis of the overall effect of wearing gloves, of glove type and fit, and of 'double-gloving' (a common practice when infection risk is known to be higher).

Glove selection

The subjects were allowed to choose the size of glove that fitted them best, with some advice from the researcher (since most had little or no experience of wearing examination gloves). The latex and vinyl gloves were comparable in dimensions for each of the five sizes, so there was no variation in best-fit glove size between the two types. The 'Larger Latex' gloves were chosen at two sizes larger than the best fit, except for the four candidates that chose the 'Large' size gloves as 'Best Fit', who were assigned 'Extra-Large' as their larger size. For the 'Double Best-Fit Latex', subjects wore two layers of 'Best Fit' latex gloves.

Location

The tests were performed in a laboratory at the University of Sheffield. Subjects were seated at a standard height table on which the test apparatus was placed.

Experimental design

The tests were performed over six different sessions on separate days, with each type of glove worn for one session, and the 'No Gloves' condition for the first and last sessions. The rationale for carrying out the tests in separate sessions were: to avoid hand fatigue, to reduce the effect of learning behavior, and to increase the availability of test subjects (each session took around 15-20 minutes, and most participants were more willing to give time in short sessions than for one session of 1.5-2 hours). It was recognized that participants' energy levels, skin moisture or other factors might vary from day to day, but it was decided that the benefits, particularly of reducing hand fatigue, outweighed the costs.

Furthermore, the order of the four gloved conditions was randomized to reduce or eliminate some of these possible confounding effects. The first and last sets of tests were performed with no gloves in order to provide a baseline measure, independent of learning behavior or glove type, to which the individual gloved tests could be compared. (For the Semmes-Weinstein Monofilaments, only one 'No Gloves' test was performed, since learning behavior is not a factor.) This also allowed for some learning to be done before the gloved conditions were tested, these being the most important for comparison. Ideally, the subjects would have performed the test multiple times before recording the results, but the available time did not allow for this. The order of the two dexterity tests was randomized to allow a fair comparison between the tests in terms of their discrimination, so that one test was not more affected by hand fatigue than the other.

Statistical analysis

Because of the 'repeated measures' nature of the experiments (i.e. the same subjects were used for each hand condition), the results were analyzed using paired difference tests. Unless otherwise stated, the significance level used is 5% ($\alpha = 0.05$), and the null hypothesis is that the difference between paired responses (i.e. two tests performed by the same subject with different hand conditions) has a mean value of zero.

Each data set was tested for normality with the Shapiro-Wilk test, which is most appropriate for small data sets (Shapiro & Wilk, 1965). For those data sets in which the null hypothesis of normality was not clearly rejected, paired t-tests were used. For those data sets in which the null hypothesis of normality was clearly rejected, the non-parametric Wilcoxon Signed Ranks Test was used. This compares the mean ranks of the samples rather than mean scores. While this means that the assumption of normal distribution of the population is not necessary, the significance of the results may be less apparent.

In order to compare the different dexterity tests fairly in terms of their ability to measure the performance differences between gloves, the mean difference in performance of each gloved condition to the 'No Gloves' condition was calculated as a percentage of the mean 'No Gloves' performance i.e. for n subjects:

$$Gloved\ Performance\ Measure = \frac{\sum_{i=1}^{n}(Gloved\ score_i - 'No\ Gloves'score_i)}{\sum_{i=1}^{n}'No\ Gloves'score_i} \times 100$$

The relative performance is shown as a bar chart for each test (e.g. Figure 4). The CSPDT is measured in terms of completion time, meaning that in order to compare 'performance' in a similar way to the Purdue, in which the number of completed insertions or assemblies is measured, the inverse of completion time must be calculated. The 95% confidence intervals in the mean percentage difference are also indicated.

For the dexterity tests, the 'No Gloves' score was generally calculated as an average of the initial and final scores (tests 1 and 6). For this to be a fair comparison with the gloved scores, which were spread across tests 2 to 5, a linear relationship between test number and score for a given hand condition must exist. To test this assumption, the mean scores for each session were plotted, and one candidate was also chosen to repeat the test a number of times in the ungloved condition. For those tests where it was considered that the assumption of linearity was not valid, the 'No Gloves' score was adjusted based on the learning curve found.

Apparatus and test procedure

Purdue Pegboard Test. The Purdue Pegboard and its administration have been described in detail elsewhere (e.g., Tiffin & Asher, 1948). It consists of a board with two columns of holes and four 'cups'. The left- and right-most cups contain metal pins, while the central two contain collars and washers. The procedure comprises four tests: Left Hand Test, Right Hand Test, Both Hands Test and Assembly.

In the first three tests, the subject is given 30 seconds to place as many pins as possible, one at a time, into the holes, with the right hand, left hand and both simultaneously, starting at the furthest hole or pair of holes and moving down the column(s). A combined score is obtained from the sum of the scores for the three tests (with the 'Both Hands' score being the number of pairs placed). In the Assembly test, the subject builds an 'assembly' in each hole using both hands alternately, starting with a pin, then placing a washer, a collar and another washer onto the upright pin. The subject is given one minute to complete as many assemblies as possible. The score is obtained from the total number of parts assembled (1 assembly = 4 parts). Since the risk of dropping instruments or materials can be very important to performance in any surgical discipline (medicine or dentistry) it was decided to record any dropped parts. Subjects were instructed to pick a new part from the cup if they dropped one, rather than attempting to retrieve the dropped one, so as not to add a further time penalty.

Crawford Small Parts Dexterity Test (CSPDT). The CSPDT, described in Crawford and Crawford (1956), consists of two parts. In the 'Pins and Collars' test, subjects use tweezers to place pins in a holed board and then place flanged collars over them (before moving on to the next pin). The test score is the time taken to complete 36 pin-collar assemblies in six rows. A practice row of six holes is provided. As with the Purdue tests, the number of dropped parts was also recorded, although this is not part of the standard scoring.

In the ‘Screws’ test, subjects pick up custom screws by hand and screw them into threaded holes until the threads have just engaged. A flat head screwdriver is then used to screw them down until the threads disengage and they drop onto the tray beneath. As with the ‘Pins and Collars’ test, the test score is the time taken to screw in 36 screws in six rows, and a practice row of six holes is provided.

Preliminary testing showed that the ‘Screws’ test was taking well over ten minutes to complete and was causing serious hand fatigue. This was much longer than would have been expected given the data provided with the test. Discussions with the supplier did not resolve the discrepancy, and so it was decided to shorten the test by asking the subjects to complete only two rows (12 screws).

Semmes-Weinstein Monofilaments. An example of the Semmes-Weinstein Monofilaments apparatus can be seen in Figure 1. It consists of a set of twenty nylon monofilaments, each perpendicularly attached to a separate plastic handle. The monofilaments are equal in length (approximately 40mm) but vary in diameter. The handles are each marked with a letter, from A to T, and a number representing the force level, which is calculated as follows:

$$Force\ level = log[buckling\ load\ (g) \times 10^4] \qquad (1)$$

The force level ranges from 1.65 (A) to 6.65 (T), which corresponds to a range of 4.38 x 10^{-5} – 4.38 N in buckling load.

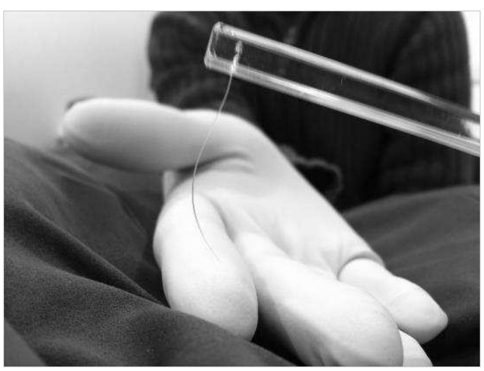

Figure 1. Semmes-Weinstein Monofilament Test

A number of issues have been identified with the current equipment and procedure. The main problem is with accuracy and repeatability of the applied force. Because of the nature of the mechanoreceptors close to the surface of the skin, which are particularly sensitive to edge effects, the level of stimulation can change depending on whether these mechanoreceptors encounter the edge of the filament. The applied force can also change dramatically depending on the friction conditions, since the end can be considered as either free or pinned, the difference theoretically changing the buckling force by a factor of 16. Furthermore, the applied force may be subject to dynamic effects, in which the buckling load can be exceeded if the force is applied too quickly.

The accuracy of the specified forces has also been questioned (J. Bell-Krotoski & Tomancik, 1987; Weinstein, 2010) since the monofilaments are often manufactured to size specifications rather than being calibrated for force, and so variation in the properties of the nylon could affect the buckling load. Bell-Krotoski and Tomancik also noted that the contact stress was almost impossible to calculate because of the bending of the filament, although it is unclear whether stress or applied force correlate best with cutaneous stimulation. Since the diameter of each filament is different, this is an important question, since the applied stress may not correlate with nominal force. However, it may be that where the diameter of the filament is smaller than the spacing of the mechanoreceptors, the difference between force and stress becomes immaterial. Further work that is beyond the scope of this study would be needed to fully understand this area.

Lastly, the nature of the monofilaments means that the applied force varies in discrete amounts, which limits the resolution of the test. Since the differences in the effects of medical gloves on tactility may be very slight, they may be difficult to identify at the current resolution.

Some attempts have been made to solve the issues mentioned – notably the introduction of the Weinstein Enhanced Sensory Test (WEST), which has rounded ends and individually force-calibrated filaments to produce a more consistent buckling load. However, the WEST has a reduced number of filaments (two handles with five filaments each), so that the resolution is more coarse than in the original test. It is claimed that the WEST has greater accuracy and repeatability, but testing has shown the Semmes-Weinstein test to be comparable with other available tests in terms of repeatability (J. Bell-Krotoski & Tomancik, 1987). Using the same filament set and operators across the range of hand conditions for each subject will further improve repeatability, and the need for finer resolution meant that the Semmes-Weinstein Monofilaments apparatus was ultimately preferred over the WEST.

The testing procedure used was the Rapid Threshold Procedure™ (from (Weinstein, 2010)), which seeks to determine the threshold force at which detection occurs fifty per cent of the time. The procedure is as follows:

1. Start well above the threshold and move down the force scale
2. Ensure the participant cannot see the filament
3. Apply the filament to the fingertip steadily (approximately one second each for application, holding and removing)
4. If the participant indicates that they detected the force, proceed to the next lowest force
5. At the first failure to detect, go back to the previous (higher force) filament and test again
6. If they fail to detect this filament, its value is the threshold (since they have once succeeded and once failed to detect it)
7. If they do detect it, move down to the previously-missed level and stimulate again
8. If they miss this level (for the second time) the threshold is taken as halfway between the higher, detected and lower, undetected values
9. If they detect this level, proceed to the next lowest level as if they had never missed it

Examples are shown in Table 1.

Table 1. Examples of the Rapid Threshold Procedure™ (Weinstein, 2010)

Level	Detected?	Level	Detected?	Level	Detected?
4.17	Yes	4.17	Yes	4.17	Yes
4.08	Yes	4.08	Yes	4.08	No
3.84	No	3.84	No	4.17	Yes
4.08	No	4.08	Yes	4.08	Yes
		3.84	No		

Threshold: 4.08   Threshold: 3.96   Continue to 3.84

RESULTS

Purdue Pegboard Test: Combined (Left, Right and Both Hands)

Learning behavior. It was important to establish the learning behavior of subjects in order to determine the fairest way to compare the ungloved and gloved results. This was done in two ways: by taking the mean score for each test session, and by having one naive subject perform the test repeatedly in the ungloved condition. The results are shown in Figure 2.

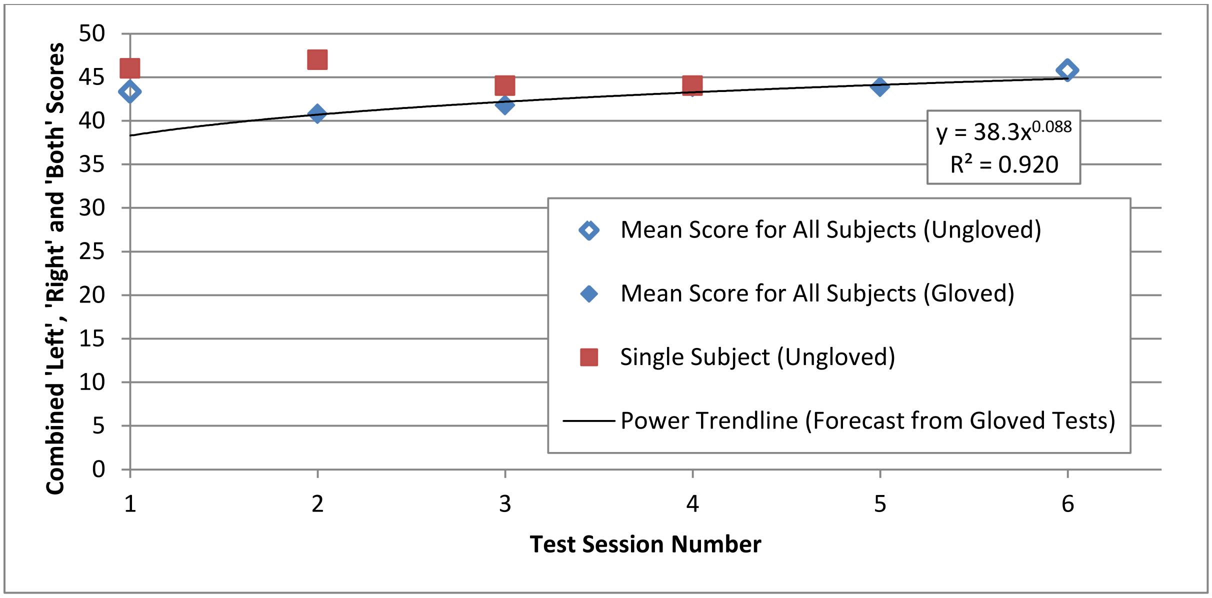

Figure 2. Learning behavior for the Purdue Pegboard (Combined)

Learning behavior is expected to be non-linear, being steepest at the beginning and leveling off after some time so that no further learning occurs. The gloved results (tests 2 to 5) fit a power trend with a correlation of 0.92, varying slightly from linearity. However, the single subject tests show a much flatter learning curve. Furthermore, while the mean ungloved scores (tests 1 and 6) are clearly higher than the forecast glove scores for those tests, the mean score for the first test is much higher above the curve than for the last test, suggesting that the learning curve is not as steep as predicted by the power trend. A linear trend for the gloved data would still give a correlation of 0.90 and assuming linearity requires less manipulation of the data. Taking an average of the two ungloved scores is therefore thought to be the best method for fair comparison.

Results. The results of the combined Left Hand Test, Right Hand Test and Both Hands Test are shown in Figure 3. It can be seen that the best score was achieved in the ungloved condition, while the worst score was achieved with the larger gloves. The vinyl scored slightly higher than the latex. The most pins were dropped when wearing a double layer of latex gloves (0.91 per test), and the least with the larger latex gloves (0.55 per test).

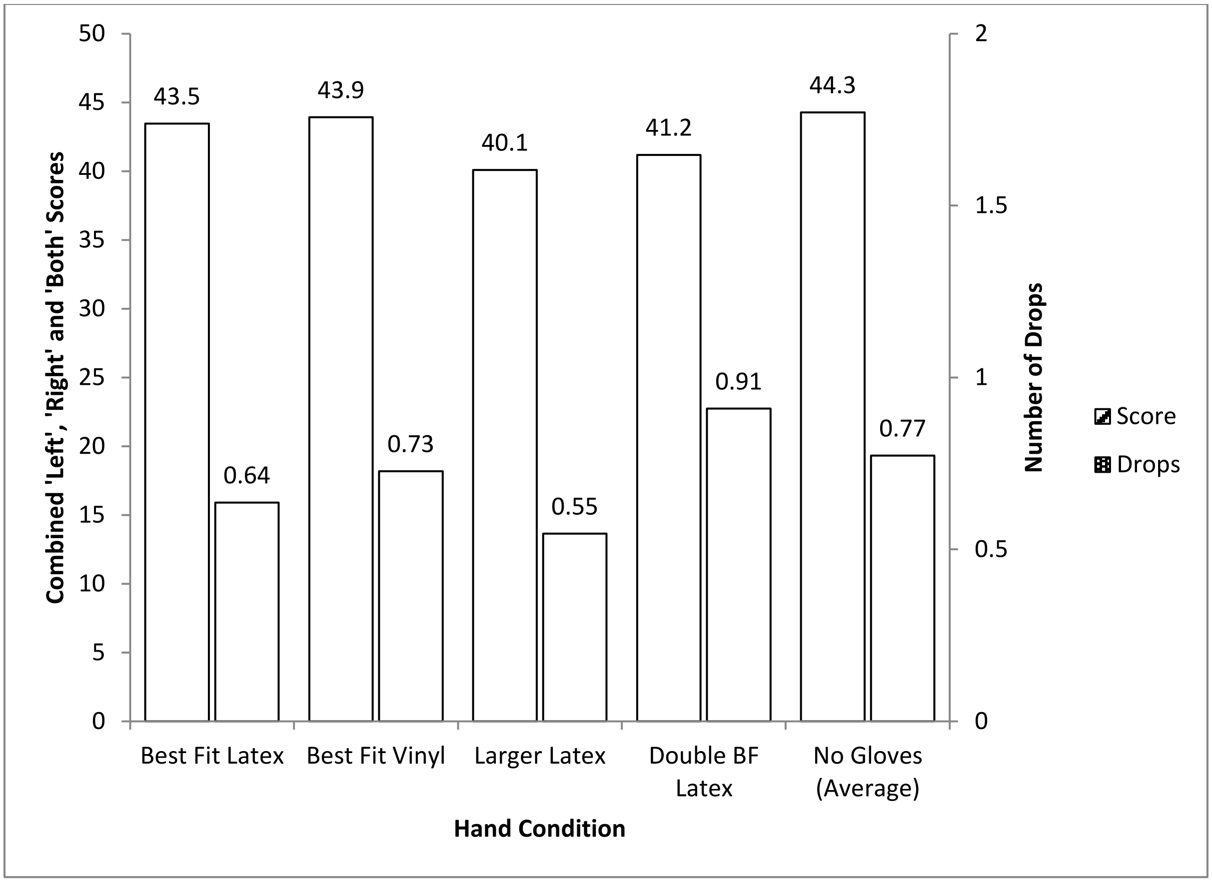

Figure 3. Mean combined scores and number of drops from the 'Left Hand', 'Right Hand' and 'Both Hands' Purdue Pegboard Tests

Figure 4 shows the relative performance of the four gloved conditions to the ungloved condition (using only the combined score, and taking the average of the two ‘No Gloves’ tests again). The 95% confidence levels are shown for an indication of significance.

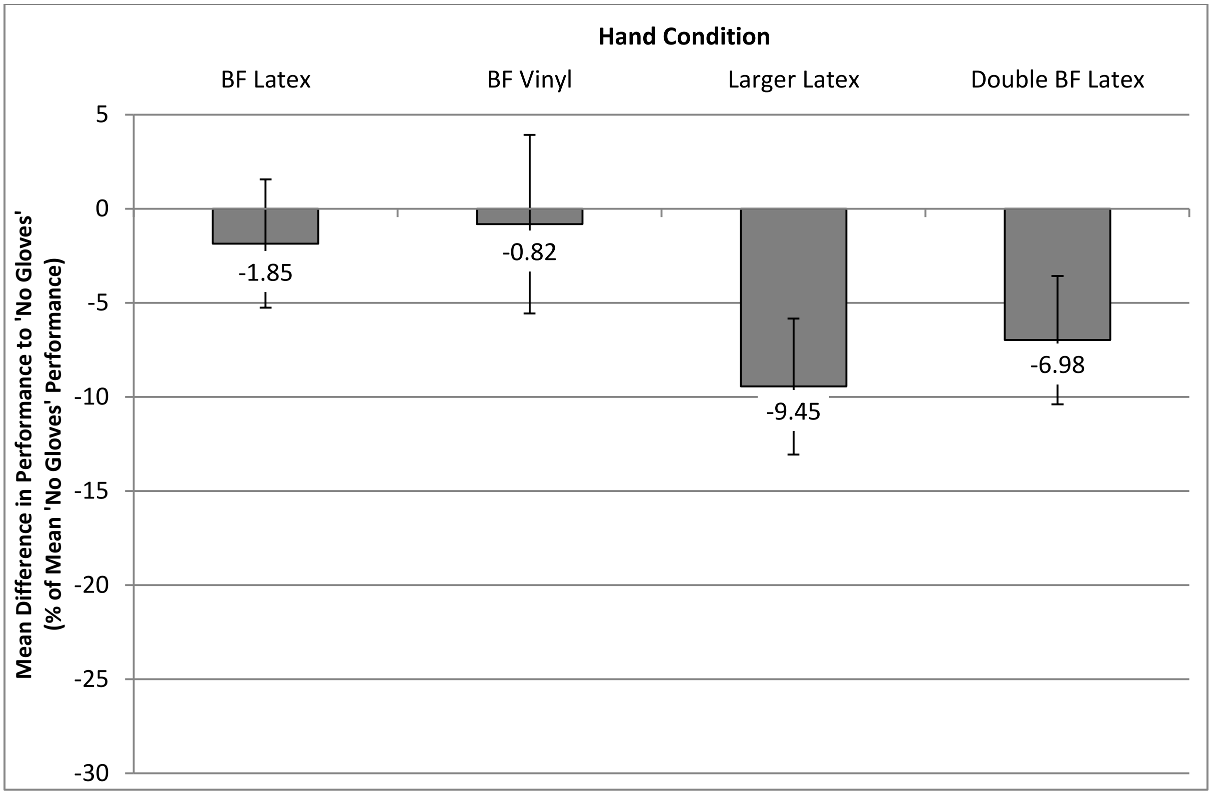

Figure 4. Comparison of Purdue combined (Right + Left + Both Hands) scores for different hand conditions with average ‘No Gloves’ score (shown as the mean difference to the ‘No Gloves’ score as a percentage of the mean ‘No Gloves’ score, with values shown below columns and 95% confidence intervals indicated)

The Shapiro-Wilk statistical test showed no significant deviation from normality for any hand conditions ($p \geq 0.264$), so repeated-measures ANOVA was used. Hand condition was found to have a significant effect on performance ($p = 0.002$). The results of paired t-tests between each of the hand conditions are shown in Table 2 (where ‘NS’ indicates no significant difference and ‘S’ indicates a significant difference between the two conditions). It was found that the results split into two groups that were significantly different from each other. The two single-layer, best-fit glove conditions and the ungloved condition were not significantly different from one another, but the larger latex gloves and the double layer of gloves produced a significantly worse performance in the combined test (Figure 5). Those that performed best are to the left of the diagram, with the worst being on the right. Variables that overlap in the horizontal axis are not significantly different from each other, while those between which a horizontal gap exists differ significantly in their performance. The diagrams are entirely schematic, and the size and spacing of the boxes are not exactly proportional to any statistical values.

Table 2. Paired t-test for Purdue (Combined) results

	Best Fit Latex	Best Fit Vinyl	Double BF Latex	Larger Latex
No Gloves	NS (0.404)	NS (0.772)	S (0.001)	S (0.000)
Best Fit Latex		NS (0.951)	S (0.010)	S (0.018)
Best Fit Vinyl			S (0.034)	S (0.031)
Double BF Latex				NS (0.802)

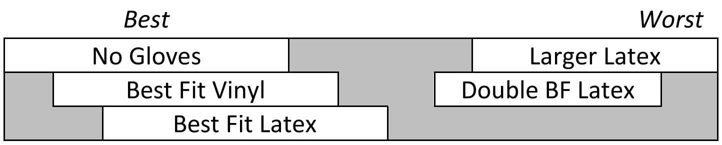

Figure 5. Schematic of significance for Purdue (Combined) results

Purdue Pegboard Test: Assembly

Learning behavior. As with the combined results, the assembly test shows a slight non-linearity (Figure 6), as expected, but the single subject tests do not show a steep learning curve, which is a somewhat unexpected result. Larger-sample testing of learning behavior would give a clearer picture. However, based on the available data, the assumption of linearity is a reasonable one.

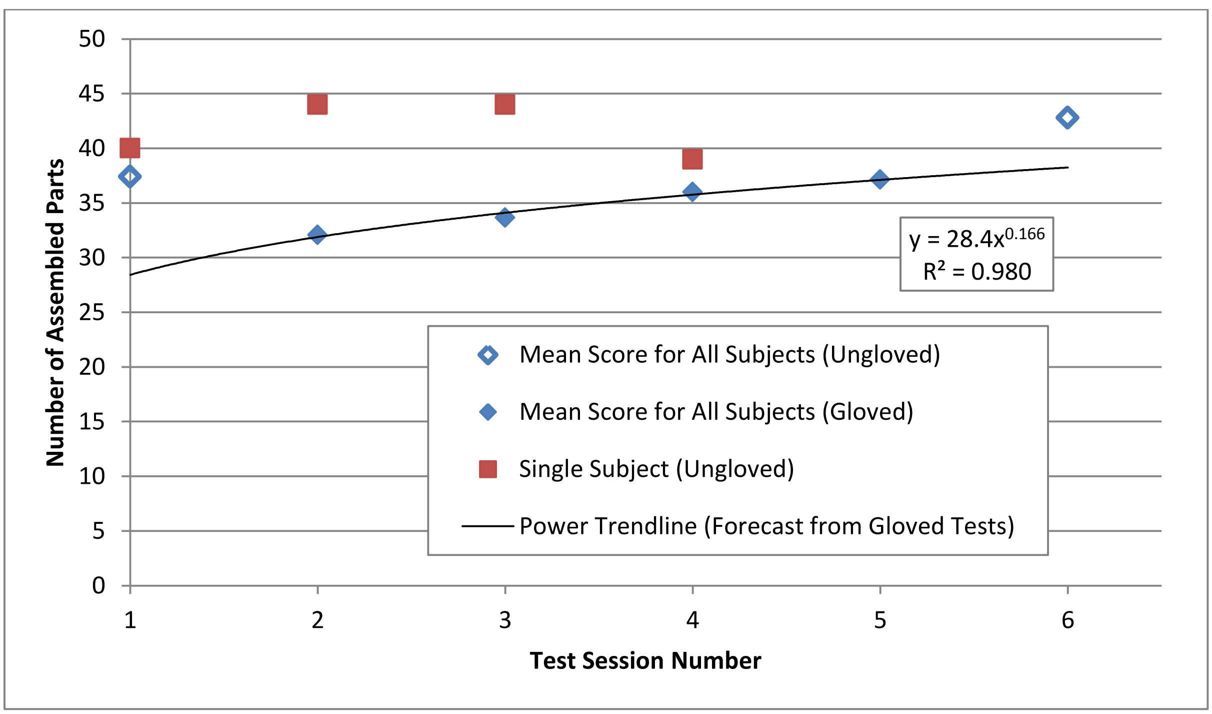

Figure 6. Learning Behavior for the Purdue Assembly Test

Results. The results of the Assembly test are shown in Figure 7. The highest score was achieved in the ungloved condition again, while the lowest score was again achieved with the larger gloves. In contrast to the combined tests, the 'Best-Fit Latex' scored higher than the 'Best-Fit Vinyl'. As with the combined tests, the most drops occurred when wearing a double layer of latex gloves, but the least occurred with the vinyl gloves.

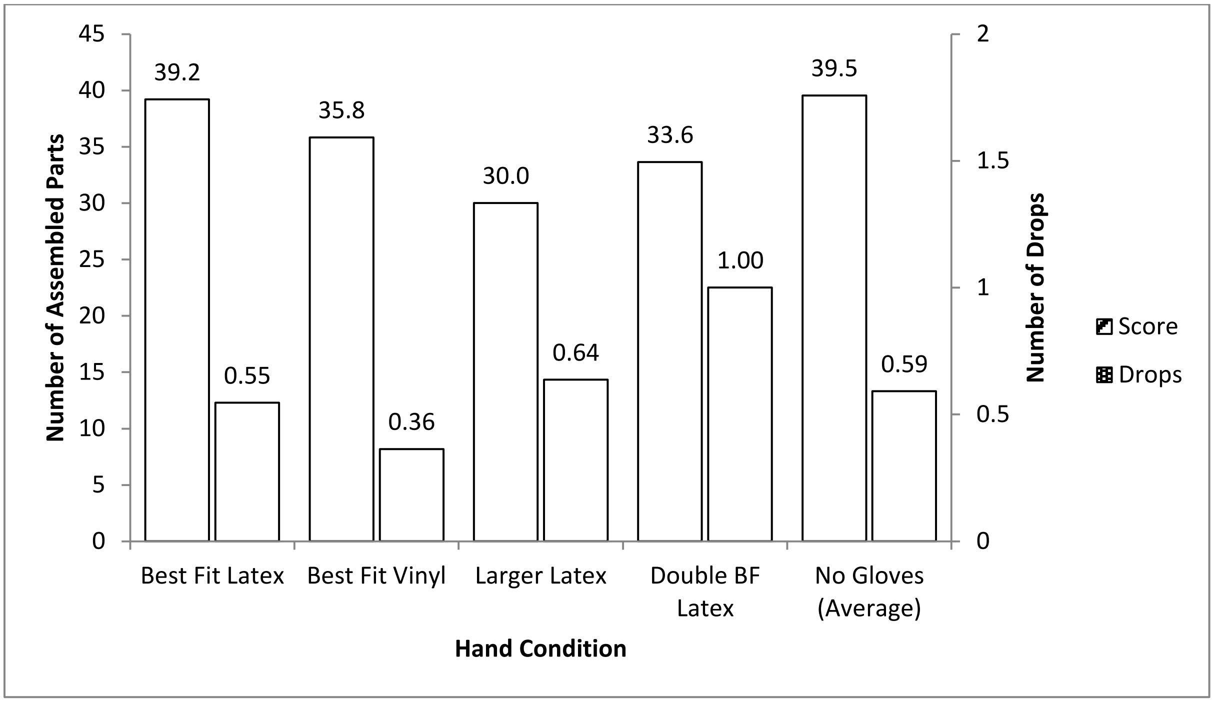

Figure 7. Mean number of assembled parts and number of drops for the Purdue Assembly test

Figure 8 shows the relative performance (number of assembled parts) of the four gloved conditions to the ungloved condition, along with 95% confidence levels.

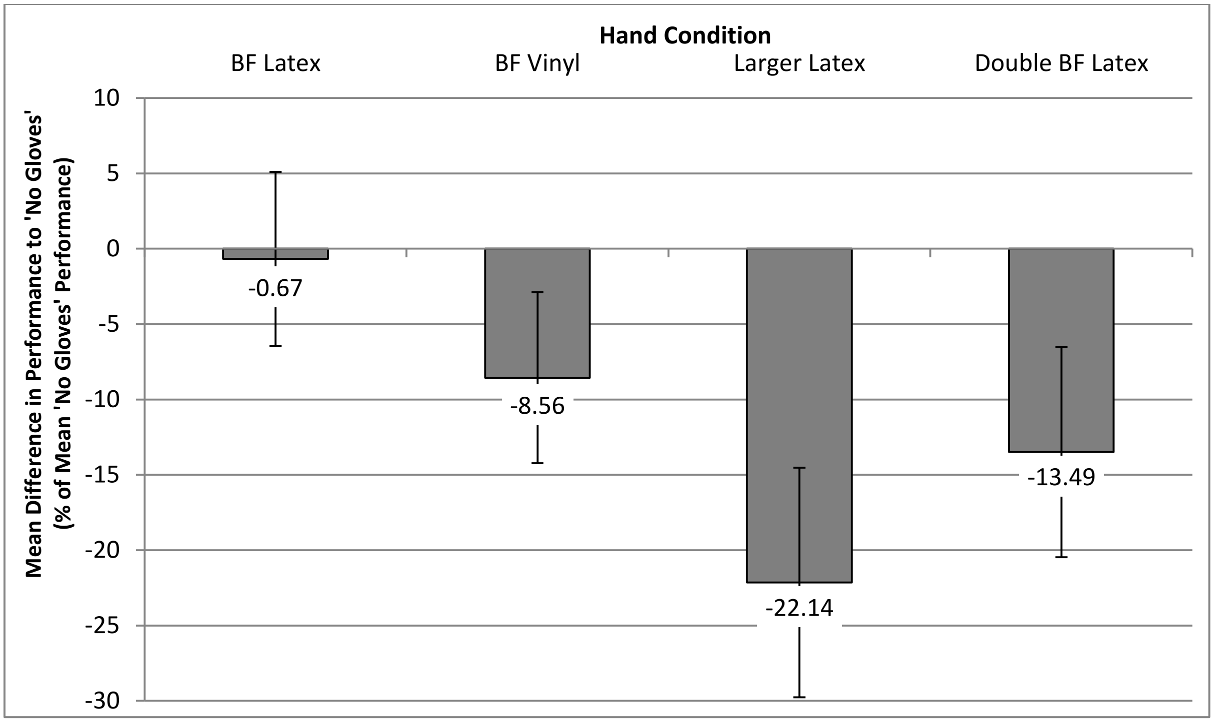

Figure 8. Comparison of Purdue Assembly scores for different hand conditions with averaged bare-handed score (with 95% confidence intervals)

The Shapiro-Wilk test showed no significant deviation from normality for four of the five hand conditions, so repeated-measures ANOVA was used. Hand condition was found to have a significant effect on performance (p = 0.000). Paired t-tests between each of the hand conditions (Table 3) found that the mean performance for the ‘No Gloves’ and ‘Best-Fit Latex’ were significantly higher than for the ‘Larger Latex’ and ‘Double Best-Fit Latex’, but the performance with vinyl gloves was not significantly different from any of the other conditions. Figure 9 is a schematic of the significance of differences between the variables.

Table 3. Paired t-test for Purdue Assembly results

	Best Fit Latex	Best Fit Vinyl	Double BF Latex	Larger Latex
No Gloves	NS (0.674)	NS (0.115)	S (0.000)	S (0.000)
Best Fit Latex		NS (0.256)	S (0.004)	S (0.012)
Best Fit Vinyl			NS (0.115)	NS (0.068)
Double BF Latex				NS (0.358)

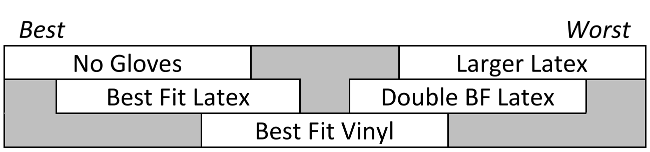

Figure 9. Schematic of significance for Purdue Assembly results

CSPDT ‘Pins and Collars’

Learning behavior. Figure 10 shows the learning behavior for the ‘Pins and Collars’ test, including the mean result across all subjects (tests 1 and 6 being ‘No Gloves’ and tests 2-5 being gloved) and the results for the one subject who completed four ungloved tests.

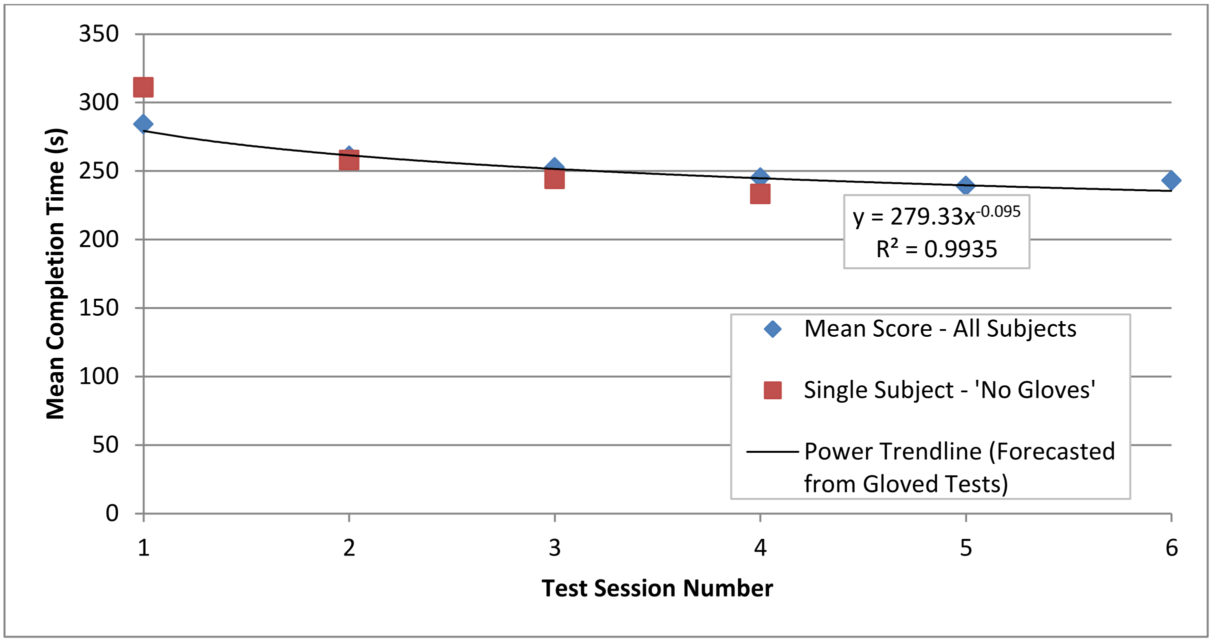

Figure 10. Learning Behavior for Crawford ‘Pins and Collars’ Test

The single-subject tests suggest a steep learning curve, and therefore taking the average of the first and last session scores may not be a fair comparison. Given the nature of a learning curve and its tendency to flatten out as learning increases, taking the final ‘No Gloves’ score may be a more reliable indicator. Using the power curve equation, the final ‘No Gloves’ score was corrected (an increase of 6% in mean time taken) to make it comparable with the gloved tests (the order of which were randomized between tests 2-5). This still made it 3.4% lower than taking the average of the two ‘No Gloves’ tests.

Results. The results are shown in Figure 11. The lowest mean completion time was achieved in the ‘Best-Fit Latex’ and ‘Best-Fit Vinyl’ conditions. The worst performance across the test group as a whole was with the double layer of latex gloves. Figure 12 shows the relative performance of the gloved conditions to the ungloved condition, performance being defined as the inverse of completion time. Only when ‘double-gloving’ did subjects perform worse on average than in the ungloved condition, but the variation in relative performance is large for all conditions. The most drops occurred when ungloved, and the least with the vinyl gloves.

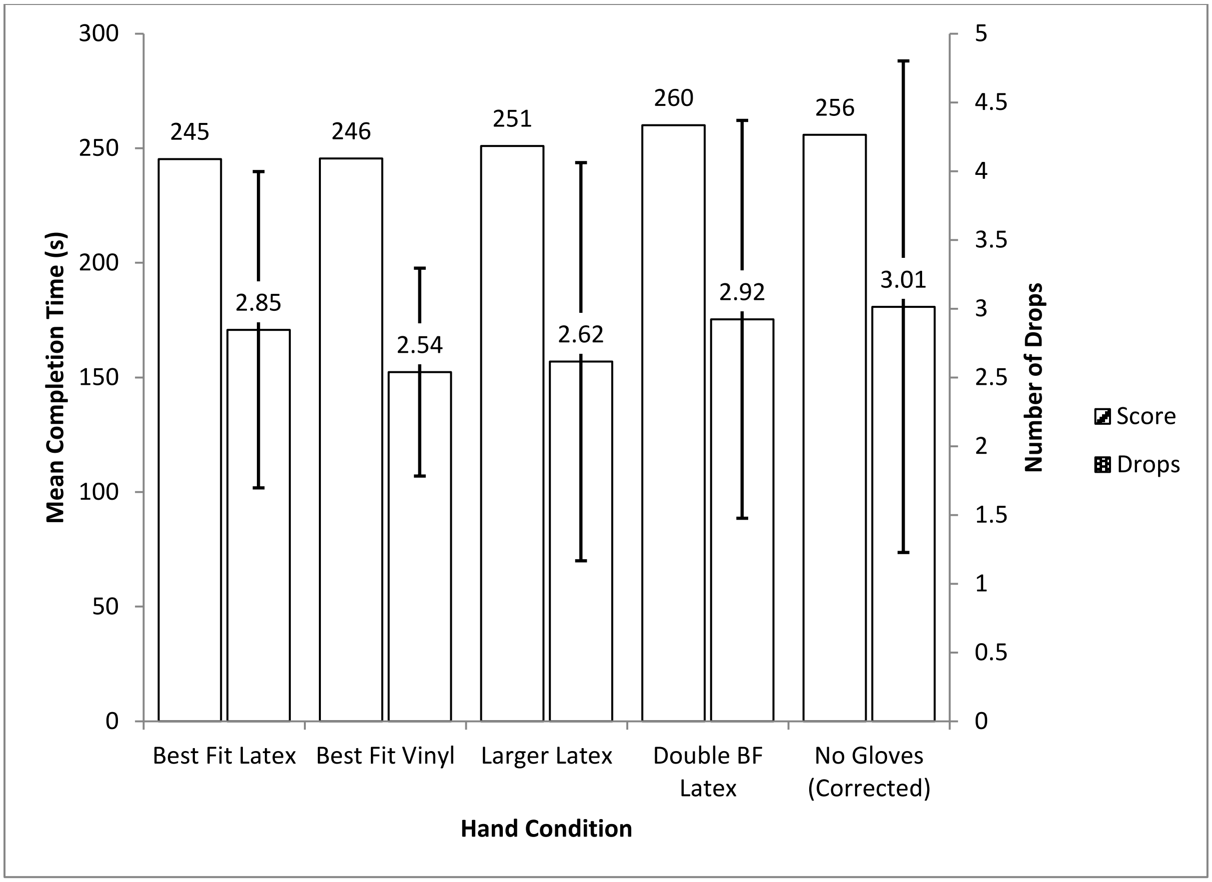

Figure 11. Comparison of Hand Conditions for Crawford – ‘Pins and Collars’ Test

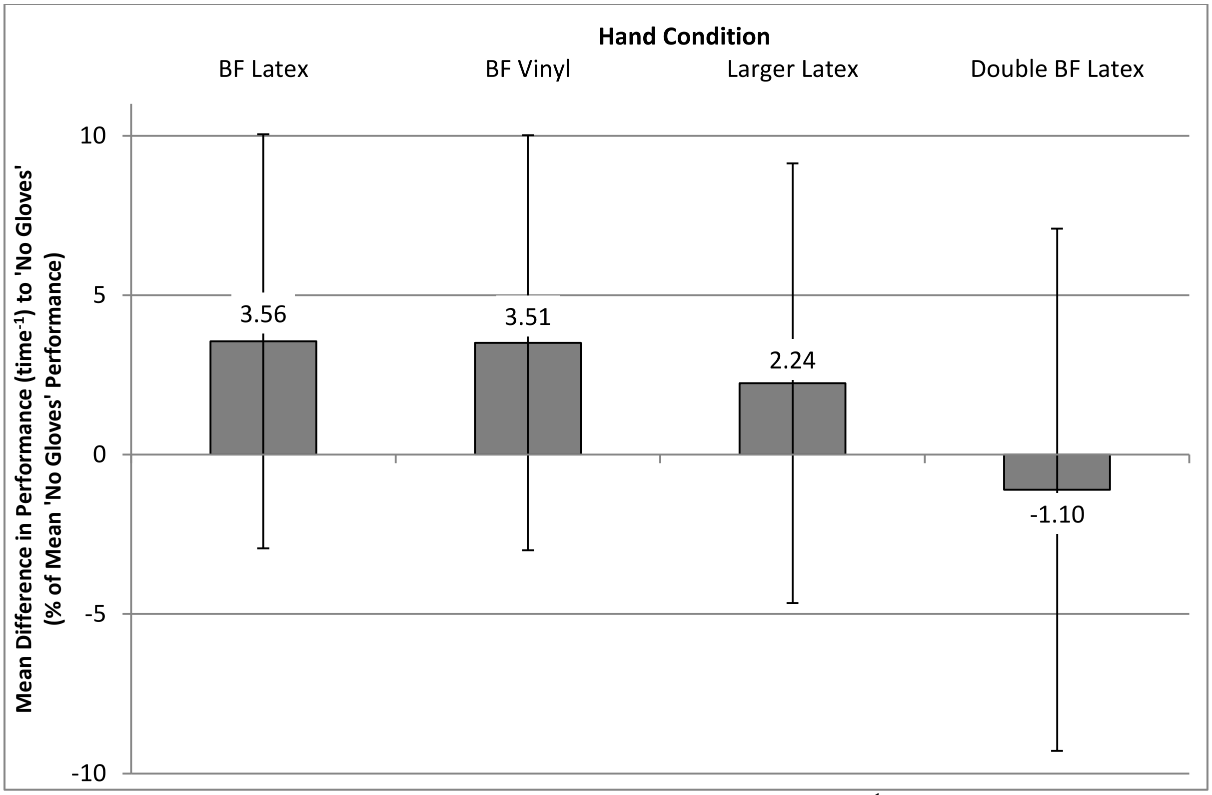

Figure 12. Comparison of Crawford – 'Pins and Collars' Test performance ($time^{-1}$) for different gloved conditions with corrected 'No Gloves' condition (with 95% confidence intervals)

None of the data showed significant non-normality ($p \geq 0.072$). Repeated-measures ANOVA for hand condition did not show any significant differences between hand conditions ($p = 0.164$). Furthermore, there were significant differences between subjects in the mean percentage of 'No Gloves' score across the four gloved conditions when compared to within-subject variation between hand conditions ($p=0.09$) i.e. the variation between subjects was more marked than the variation between hand conditions. Hand condition was also not a significant factor in the number of dropped parts in each test ($p=0.703$). In other words, the results found no consistent effect of hand condition on performance.

CSPDT 'Screws'

Learning behavior. There is a weak correlation ($R^2 = 0.46$) in the learning curve (Figure 13), but both the mean scores of the gloved tests and the single-subject ungloved tests show a reduction in completion time with repetition of the test. Since the extent of learning was unclear and a linear relationship could not be assumed, it was decided that the final 'No Gloves' score was a more reliable indicator of performance. As before, it was scaled to the average of tests 2-5 using the power curve equation (an increase of 7.5%, giving a mean completion time 4.2% less than taking the average of the two ‘No Gloves’ tests).

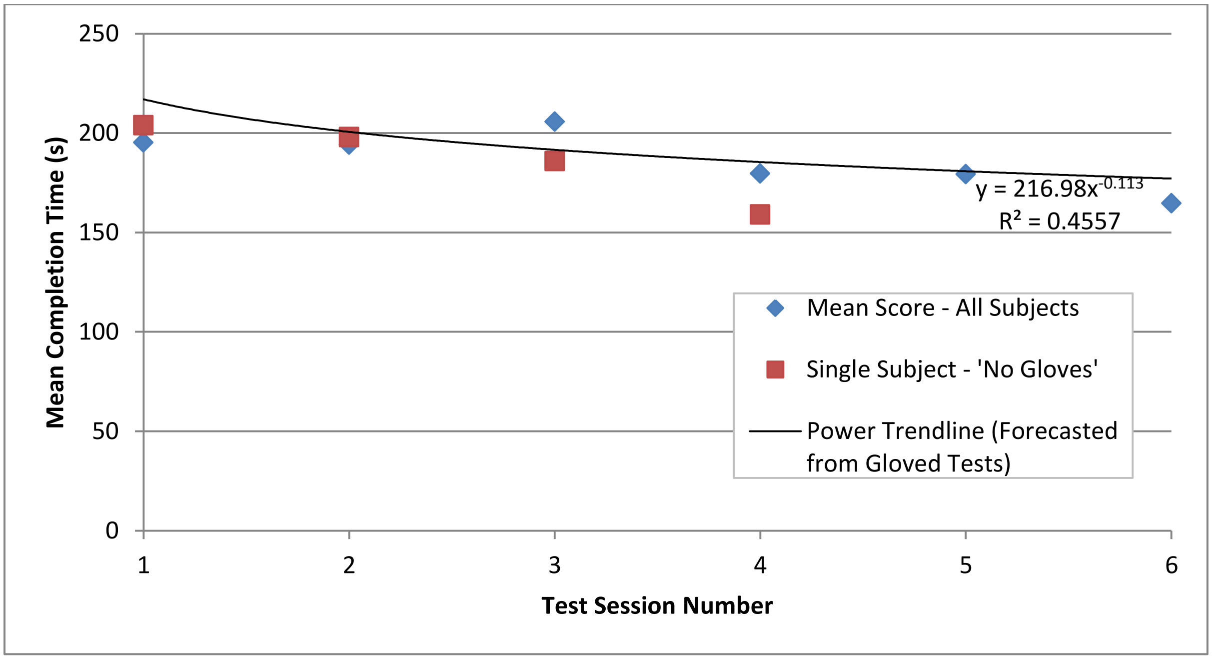

Figure 13. Learning behavior for CSPDT 'Screws' test

Results. The results of the experiment are shown in Figure 14. The shortest mean completion time was achieved with the vinyl gloves, with the longest occurring with the double layer of latex gloves. The number of dropped parts followed a similar pattern, although more drops occurred in the ungloved condition than in any of the gloved conditions. The relative performance of the gloved conditions is compared in Figure 15.

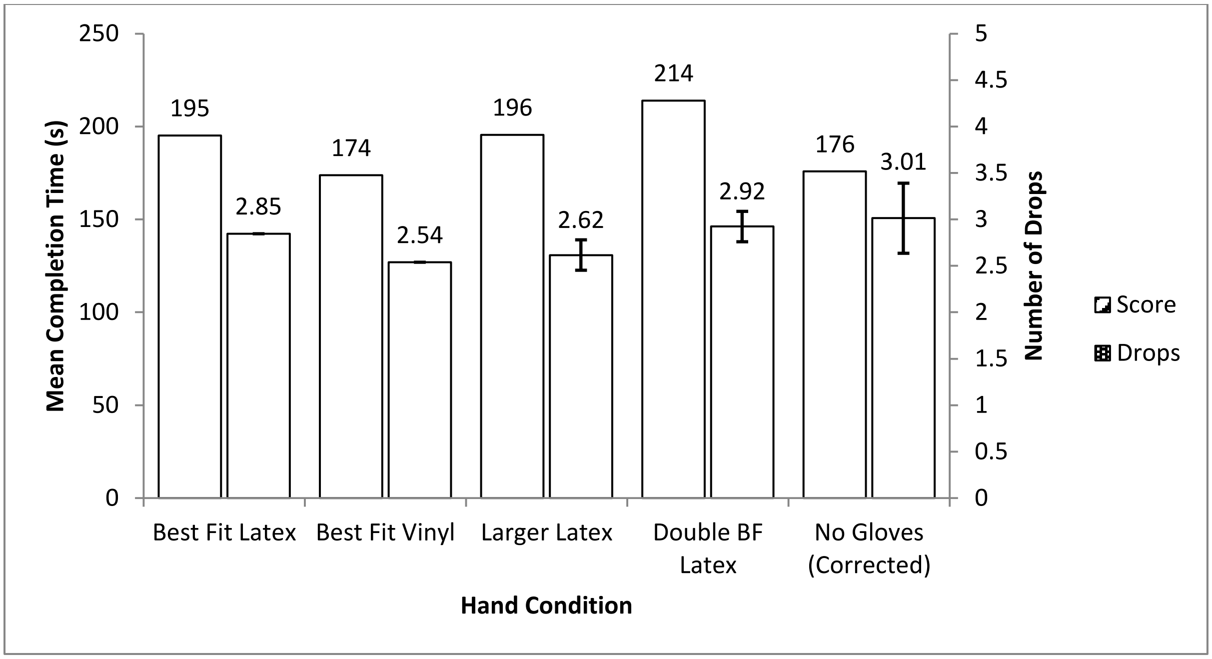

Figure 14. Mean completion time and number of dropped parts for five hand conditions in the CSPDT 'Screws' test (including 95% confidence intervals for the number of drops)

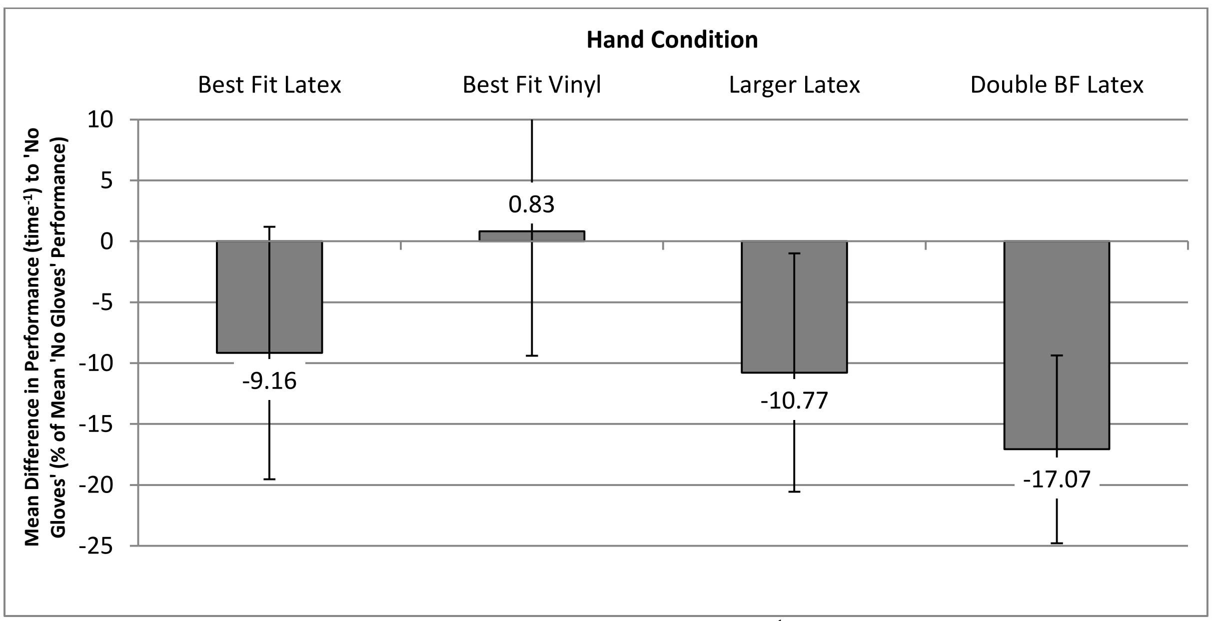

Figure 15. Comparison of Crawford – 'Screws' Test performance ($time^{-1}$) for different gloved conditions with corrected final 'No Gloves' condition (with 95% confidence intervals)

Hand condition clearly has some significant effect on performance time. Repeated-measures ANOVA (no results showed significant non-normality, $p \geq 0.166$) confirms that there are significant differences between the means (p=0.001). The results of the paired t-tests between the hand conditions are shown in Table 4, and Figure 16 is a schematic of the differences (N.B. due to overlapping, not all relationships could be displayed correctly).

Table 4. Paired t-tests for CSPDT 'Screws'

	Best Fit Latex	Best Fit Vinyl	Larger Latex	Double BF Latex
No Gloves	NS (0.058)	NS (0.900)	S (0.043)	S (0.001)
Best Fit Latex		S (0.009)	NS (0.896)	NS (0.173)
Best Fit Vinyl			NS (0.052)	S (0.004)
Larger Latex				NS (0.082)

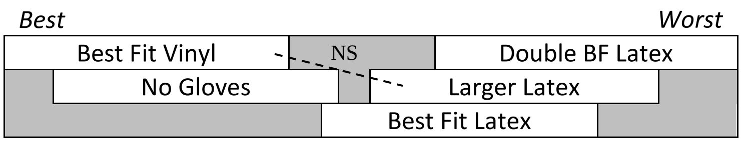

Figure 16. Schematic of differences between hand conditions for CSPDT 'Screws'

The 'Best-Fit Vinyl' and 'No Gloves' conditions clearly produce the best performance, the paired differences of both with the bottom three conditions having p values of less than 0.06 (four of six pairs being below the 0.05 significance level).

Semmes-Weinstein Monofilaments

The threshold force level varied from 2.005 (which is classified as a 'normal' level of sensation on the standard scale; Judith Bell-Krotoski, Weinstein, & Weinstein, 1993) to 4.125 (which indicates 'diminished protective sensation' i.e. a reduction in ability to feel stimuli which may be causing injury). The mean threshold forces for the five hand conditions are shown in Figure 17. Because of time restrictions on the testing, only 10 participants performed the test with all five hand conditions, with another 8 being tested with some hand conditions. Only the results of the 10 participants who completed the test in all conditions are included in the mean scores, but the results of those who completed some tests were used to calculate the paired differences where data existed.

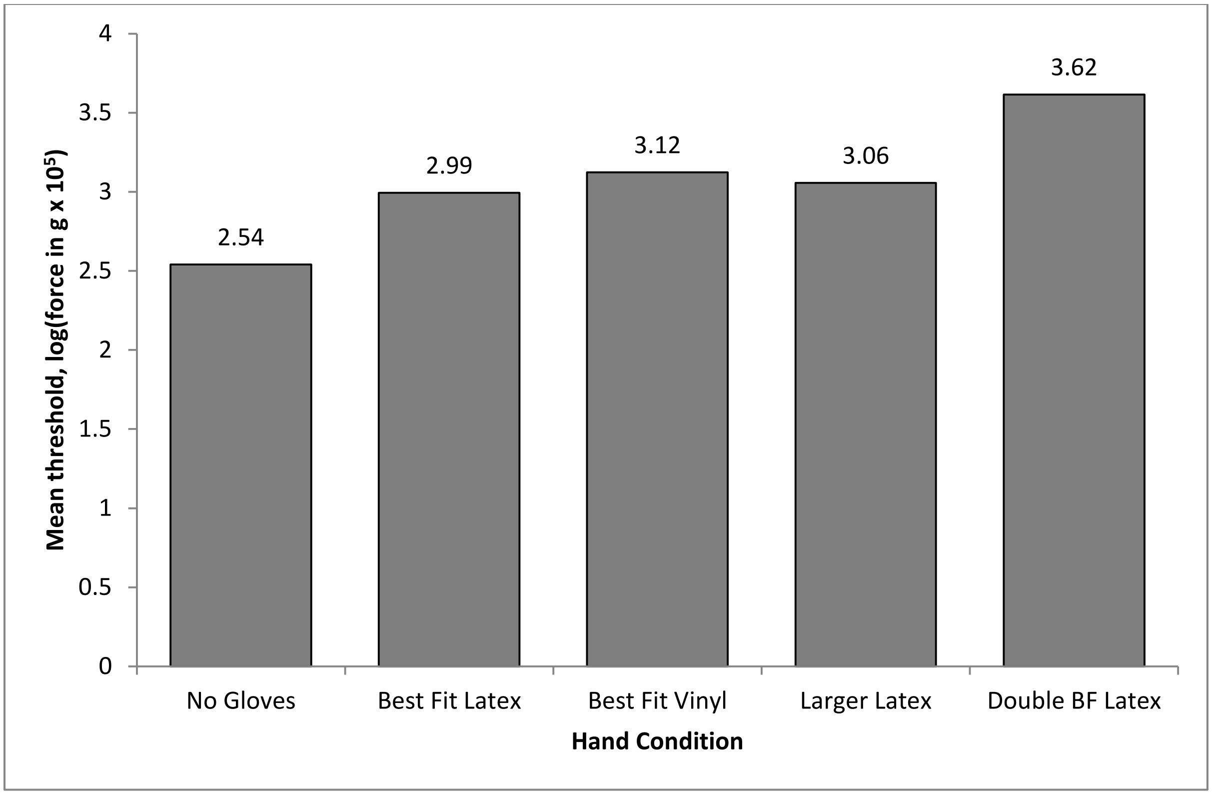

Figure 17. Mean Semmes-Weinstein Monofilament test scores (threshold log force) for five hand conditions (n=10)

It can be seen that the lowest mean threshold force, and therefore the best performance, was achieved in the ungloved condition, with the highest mean force being achieved with the double layer of gloves. Figure 18 shows the performance (defined as the inverse of threshold force level) of the four gloved conditions relative to the ungloved condition, with 95% confidence intervals.

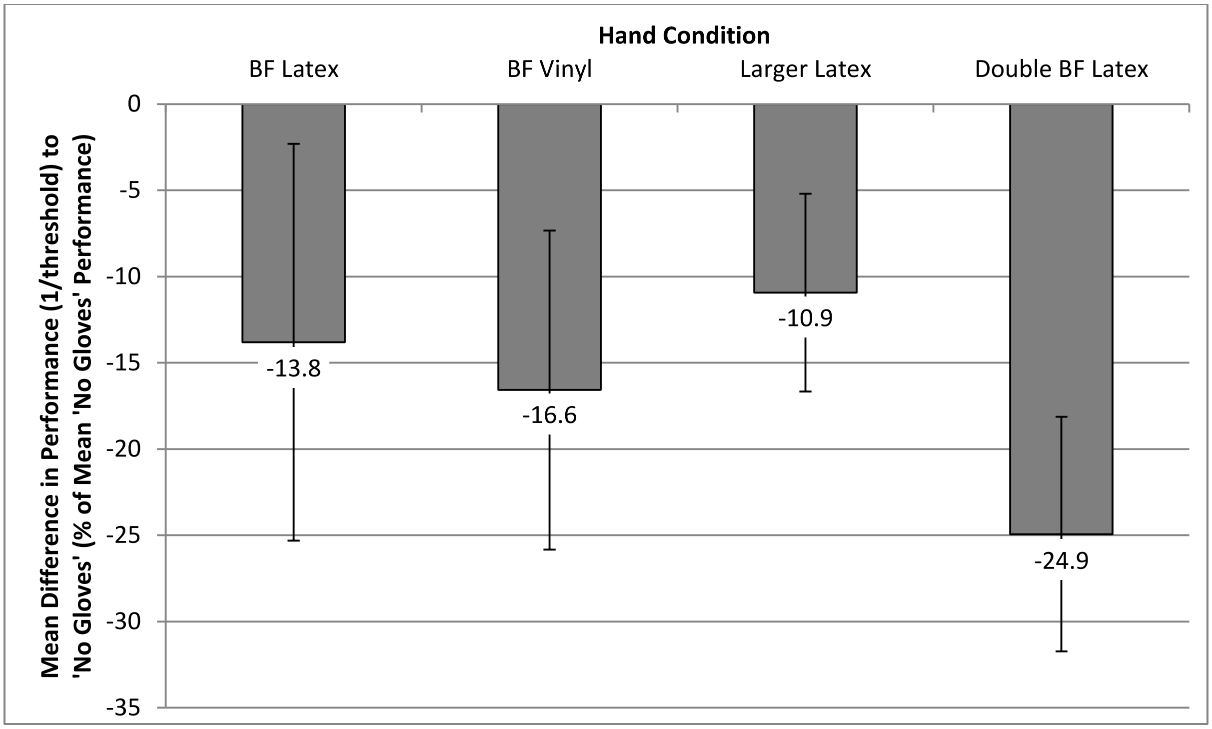

Figure 18. Comparison of differences to 'No Gloves' performance in Semmes-Weinstein Monofilament test for four examination glove conditions (with 95% confidence intervals)

Two of the results showed significant deviation from normality in the Shapiro-Wilk test. The Friedman non-parametric test for significance was therefore used. Hand condition was found to have a significant effect on threshold force (p=0.000). The Wilcoxon Signed Ranks Test was performed on pairs of hand conditions, and the results are shown in Table 5 and Figure 19.

Table 5. Paired tests (Wilcoxon) for Semmes-Weinstein monofilament test results

	Best Fit Latex	Best Fit Vinyl	Larger Latex	Double BF Latex
No Gloves	S (0.006)	S (0.003)	S (0.003)	S (0.001)
Best Fit Latex		NS (0.553)	NS (0.964)	S (0.006)
Best Fit Vinyl			NS (0.823)	S (0.016)
Larger Latex				S (0.001)

Figure 19. Schematic of significance for Semmes-Weinstein monofilaments results

The mean threshold force for the ungloved condition was significantly lower than for all gloved conditions. The mean force threshold for the double-gloved condition was significantly worse than for all the single-layer, gloved conditions. However, differences in glove size and material did not produce significant differences in threshold force.

DISCUSSION

Purdue Pegboard Test

Both the Purdue Combined (Left + Right + Both) and the Purdue Assembly tests showed similar trends in score, with the highest score being achieved with ungloved hands, followed by the two best-fit single-layer examination gloves, the double layer of best-fit latex gloves, and the larger latex gloves performing worst. Some subjects commented that the loose material of the larger gloves tended to catch, particularly on the smaller parts of the assembly, but also in the holes, making it difficult to release the pins, which could account for the performance reduction.

Wearing a double layer of gloves also significantly reduced performance compared to a single layer of the same gloves in both tests, and also produced by far the most drops in both tests. A possible reason for this increase in drops is the reduction in cutaneous sensibility caused by the extra layers. The extra thickness reduces the ability of tactile cues to be felt, and the movement of the two layers against each other could be distorting the signals further. This tactile feedback allows the subject to detect when parts are slipping and increase the grasping force, so a reduction in cutaneous sensibility could increase the frequency of drops.

Neither of the tests was able to find any significant difference between the two glove types (latex and vinyl) or between the ungloved and single-layer best-fit gloved conditions, although the difference was more pronounced in the assembly test. The resolution of the tests is fairly coarse. In both tests, the difference between the means of the 'No Gloves' and the 'Best Fit Latex' conditions was less than the resolution of the test (one pin, pair of pins or assembled part). A subject who successfully places one more pin could increase their score by two to three per cent compared to one who does not quite place the pin or part in time.

The significance of the differences could be increased by using a larger sample size, but this is not very practical for the amount of testing required in glove development. The resolution could be improved by increasing the test time, but this starts to introduce an element of fatigue (which was already noted as an issue for some of the participants), as well as making larger-scale testing even more difficult.

The Purdue test was originally designed to test the ability of the participants: their hand-eye co-ordination, bi-manual dexterity and brain function (such as moving both hands in order on the assembly task). These aspects are not relevant to glove design. The difference between left- and right-handed performance is also irrelevant. The Combined test is therefore not thought to show anything useful that the assembly test does not show. The test time could therefore be significantly reduced, allowing for a larger population to be tested. More pure dexterity tests, where less brain activity occurs, might find greater differences between gloves.

Crawford Small Parts Dexterity Test

Of the two parts of the CSPDT, the 'Screws' test clearly discriminates best between hand conditions. No significant differences were found in two separate experiments in the 'Pins and Collars' test between ungloved and gloved conditions with various materials, fits and even with a double layer.

Participants' comments after the 'Pins and Collars' tests were also mixed. Some found certain gloves to increase friction over the ungloved condition, while others found that the same glove reduced friction, and they disagreed on whether high or low friction was better for performance. Some commented that larger gloves made the task more difficult, while some said they made no difference. Similarly, some preferred the 'No Gloves' condition, with sweat generation in gloves being one explanation, while others felt that sweaty fingers reduced performance compared to dry gloves. Skin moisture content is clearly a factor in ungloved performance, and could contribute to the large variation between subjects. This could be improved with exfoliating and washing of hands before the test.

Latex and vinyl single-layer best-fit gloves performed very similarly (less than 0.1% difference), even though some felt the tweezers slipped more with vinyl and did not like the feel. Gripping pins with tweezers does not require large frictional forces or sliding motion, so frictional properties are probably less relevant than in other tests.

The double best-fit latex gloves did perform worst, however, and this is supported by comments made by a number of participants that double-gloving restricted movement, increased fatigue and made it harder to feel or to control muscles.

In the 'Screws' test, the vinyl gloves performed best and were significantly better than the single and double-layer latex gloves. On average, subjects completed the test two seconds faster with vinyl gloves than with no gloves, although the difference was not statistically significant. This result is surprising, but may show the importance of friction in the task, and that high friction is not always desirable. Many participants found the screwing part easier with vinyl because of the lower friction allowing the screwdriver to rotate whereas the latex gloves sometimes stuck to the screwdriver or to the screws when releasing them, but it was felt that the poor conformability of the vinyl made manipulation of the screw more difficult and the loose material sometimes caught when turning.

The same was true for both the double layer and the larger gloves, which performed worse, with loose material getting caught in the threads or on the screwdriver, meaning the technique needed to be adjusted. As previously discussed, this might not affect performance but may increase stress or discomfort. Both larger and double gloves reduced the perceived ability to manipulate the screws, whether due to loss of sensation, loose material or slipping of the two layers. Time was also wasted with the larger gloves in having to pull up the fingers to keep the glove material tight on the fingertips.

Analysis: Dexterity

Evaluation of the tests. The statistical significance of differences between the two single-layer, best-fit gloved conditions for each dexterity measure is shown in , along with the significance of the overall effect of hand condition (for n = 18 subjects and k = 5 hand conditions). Statistically significant differences are shown in bold. Only the CSPDT ‘Screws’ test found a significant difference between the glove materials in comparable conditions. The ‘Pins and Collars’ part of the test found no significant effect of hand condition, despite testing double-layered and loose-fitting gloves. The Purdue tests both found that hand condition had a significant effect, but the performance difference between the gloves was clearer with the ‘Assembly’ test than the ‘Combined’ test score.

Table 6. Statistical significance of paired differences between latex and vinyl performance in four dexterity measures and analysis of variance (ANOVA) for effect of hand condition (n = 18; k = 5)

Test	p value for Paired Difference of Best-Fit Gloves (Latex – Vinyl)	p value (ANOVA) for Effect of Hand Condition
Purdue Combined	0.951	**0.002**
Purdue Assembly	0.256	**0.000**
CSPDT ‘Pins and Collars’	-	0.164
CSPDT ‘Screws’	**0.009**	**0.001**

Effect of medical gloves on dexterity. There was no consistent effect of wearing gloves or of glove material on dexterity across the tests. The ranking of hand conditions in terms of mean performance in each test is shown in . In two of the tests, ungloved performance was the best, in one it performed worse than all best-fit gloved conditions, and in one it was better than latex, but worse than vinyl. In terms of glove material, latex ranked higher than vinyl in two tests, and lower in two others. However, performance with the double layer of latex gloves was worse than with a single layer in all tests. The same is true for larger gloves, compared to those selected as the best fit.

Table 7. Ranking of hand conditions in dexterity tests (BF = Best-Fit)

	Purdue (Comb.)	**Purdue (Assem.)**	**CSPDT (P&C)**	**CSPDT (Screws)**
1	No Gloves	No Gloves	BF Latex	BF Vinyl
2	BF Vinyl	BF Latex	BF Vinyl	No Gloves
3	BF Latex	BF Vinyl	Larger Latex	BF Latex
4	2 x BF Latex	2 x BF Latex	No Gloves	Larger Latex
5	Larger Latex	Larger Latex	2 x BF Latex	2 x BF Latex

These results fit with what is already known about glove effects from previous research (Gnaneswaran, Mudhunuri, & Bishu, 2008; Shih, Vasarhelyi, Dubrowski, & Carnahan, 2001), as well as with comments made during the testing. An added layer reduces and distorts tactile feedback signals that are vital for fine dexterity, while loose material similarly reduces tactility, but also gets trapped, causing delays in performing fine manual tasks.

The varying performances of glove materials and the ungloved condition in the tests may simply be due to statistical variation, since no significant differences were found for most of the tests, or it may be that different aspects of dexterity require different attributes. The two tests in which vinyl outperforms latex may benefit from lower friction, allowing pins to be released or a screwdriver to be turned more easily. The Purdue Assembly test requires fine finger dexterity and good tactile feedback, and so the ungloved condition performs best, and the close-fitting latex outperforms the less-elastic vinyl. In order to draw more firm conclusions about glove effects on dexterity, it is necessary to understand other factors such as their effect on tactility, their mechanical properties and the effect of glove fit.

Semmes-Weinstein Monofilaments

The relative performance of the ungloved condition and the gloved conditions is unsurprising. The gloves create a barrier between the stimulus (the monofilament tip) and the mechanoreceptors that sense mechanical stimulation.

The soft polymer of the gloves is likely to dampen the impact and to spread the load across a larger area. Increasing the thickness of the barrier (such as by double-gloving) will increase the damping.

It would also be expected that the loose material in larger gloves would distort the tactile cues and make detection of a stimulus more difficult. However, the results do not support this hypothesis. One possible explanation is that, with loose gloves, it was often necessary to pull the glove finger taut in order for the monofilament not to slip. It is recognized that this is not necessarily a realistic representation of clinical performance, but was necessary for the completion of the test. In medical practice, there will be significant movement and rucking of the gloves, which is likely to inhibit sensibility.

The lack of significant differences between gloves of different materials contrasts with medical practitioners' views (Mylon, et al., 2013) that less flexible and conformable gloves reduce tactile ability. The mean number of force levels by which the two conditions differed was 0.3, and the resolution of the test is 0.5. (Where one level is detected both times and the level below is not detected at all, the threshold is taken as halfway in-between.) This presents some difficulty in finding a significant difference, since the expected threshold level will be, as often as not, the same for both conditions. A continuous scale, as proposed above, given a good level of accuracy, could provide a way to detect these finer differences.

Although the Semmes-Weinstein monofilaments test shows some ability to discriminate between hand conditions, the resolution needs improving for medical glove testing, and it would need to be supplemented with other tactility tests that create a more realistic tactile environment, recreating the real movement of gloves on the fingertips.

Lambert, Mallos, & Zagami (2009) proposed a solenoid-operated device that allows automated, repeatable force application of the filaments to a consistent location, but does not address the issue of resolution. Another device has been patented (Low, Richardson, & Wright, 1972) in which the force can be adjusted on a continuous scale, but must still be applied manually. A new design that combines these ideas to produce a rig in which a variable force could be applied mechanically to a consistently located fingertip, using a single nylon monofilament, should be considered for development in order to achieve greater repeatability and finer resolution.

CONCLUSIONS

Both parts of the Purdue Pegboard Test found a significant effect of hand condition on performance, but no significant differences could be found between glove types (latex and vinyl examination gloves). Because of the agreement of the results and the similarity in methods between the two parts, the 'Combined' test was felt to be redundant and it was recommended that only the 'Assembly' test be used for glove evaluation. Because of the relative coarseness of the test, testing on a larger sample size was recommended to increase the significance of the differences. No significant effect of hand condition on the Crawford Small-Parts Dexterity Test 'Pins and Collars' score was found, but the 'Screws' part of the test produced significant differences between hand conditions, and showed a promising ability to discriminate between gloves. Further validation of the Purdue 'Assembly' test and the Crawford 'Screws' test for medical glove performance evaluation with a larger sample size is recommended.

In the Semmes-Weinstein Monofilaments Test, tactility was found to be significantly better in the ungloved condition than when gloved, but no significant differences in performance between glove types were found. A reduction in tactility was evident when "double-gloving" with latex gloves compared to a single layer. Development of the method and the design of new, applied tactility tests were recommended in order to fully characterize the effect of medical gloves on tactile sensitivity.

REFERENCES

Bell-Krotoski, J., & Tomancik, E. (1987). The repeatability of testing with Semmes-Weinstein monofilaments. Journal of Hand Surgery, 12(1), 155-161.

Bell-Krotoski, J., Weinstein, S., & Weinstein, C. (1993). Testing Sensibility, Including Touch-Pressure, Two-point Discrimination, Point Localization, and Vibration. Journal of Hand Therapy, 6(2), 114-123. doi: 10.1016/s0894-1130(12)80292-4

Crawford, J. E., & Crawford, D. M. (1956). Small Parts Dexterity Test. New York: Psychological Corporation.

Gnaneswaran, V., Mudhunuri, B., & Bishu, R. R. (2008). A study of latex and vinyl gloves: Performance versus allergy protection properties. International Journal of Industrial Ergonomics, 38(2), 171-181.

Griffin, D. R. (1944). Manual dexterity of men wearing gloves and mittens (Fatigue Laboratory, Trans.). Cambridge, MA: Harvard University.

Johnson, R. L., Smith, H. M., Duncan, C. M., Torsher, L. C., Schroeder, D. R., & Hebl, J. R. (2013). Factors that influence the selection of sterile glove brand: a randomized controlled trial evaluating the performance and cost of gloves. Canadian Journal of Anesthesia, 1-9.

Lambert, G. A., Mallos, G., & Zagami, A. S. (2009). Von Frey's hairs – a review of their technology and use – a novel automated von Frey device for improved testing for hyperalgesia. Journal of Neuroscience Methods, 177(2), 420-426. doi: 10.1016/j.jneumeth.2008.10.033

Low, G. M., Richardson, R. W., & Wright, D. B. (1972). US Patent No. 3,662,774. U.S. Patent and Trademark Office.

Mylon, P., Carré, M. J., Lewis, R., & Martin, N. (2011). Review of test methods relevant to medical glove design. Paper presented at the Proceedings of the 55th Human Factors and Ergonomics Society Annual Meeting, Las Vegas, NV. http://www.scopus.com/inward/record.url?eid=2-s2.0-81855224726&partnerID=40&md5=43acf20540bb96a737fee0128760958e

Mylon, P., Lewis, R., Carré, M. J., Martin, N., & Brown, S. R. (2013). *A study of clinicians' views on medical* gloves and their effect on manual performance. Manuscript submitted for publication.

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). Biometrika, 52(3/4), 591-611. doi: 10.2307/2333709

Shih, R. H., Vasarhelyi, E. M., Dubrowski, A., & Carnahan, H. (2001). The effects of latex gloves on the kinetics of grasping. International Journal of Industrial Ergonomics, 28(5), 265-273.

Tiffin, J., & Asher, E. J. (1948). The Purdue Pegboard: norms and studies of reliability and validity. Journal of Applied Psychology, 32(3), 234-247. doi: DOI: 10.1037/h0061266

Weinstein, C. (2010). WEST-hand™: Terse Manual in English Retrieved 8th December, 2011, from https://sites.google.com/site/westmanuals/westhand