

# Assessment of Containment Equipment Using a Novel Test to Determine the Effect of Containment on Laboratory Accuracy

Joanna Sawyer and Allan Bennett

Health Protection Agency, Centre for Emergency Preparedness and Response, Porton Down, Wiltshire, United Kingdom

## Abstract

*A newly developed test of operator manual dexterity in containment equipment more closely replicates standard laboratory techniques than previously used commercial pegboard dexterity tests. The test measures pipetting accuracy and has been developed and fully validated to assess the effect of containment equipment on laboratory techniques. Intra-assay variation had a mean of 1.4%; inter-assay variation had a mean of 3% within day and 3.6% between days. The test retest coefficient, using Pearson's correlation, was 0.898, ( $p=0.001$ ). The test was then used to measure pipetting accuracy in class II and III biological safety cabinets (BSCs). When experienced users took the tests in the two types of BSCs, their accuracy did not differ significantly. External factors (age, experience, glove size, gender) did not influence technical accuracy within the BSCs. The participants' level of pipetting accuracy is an inherent ability not influenced significantly by the factors measured, including the type of BSC used. Operators did, however, take significantly longer to complete tasks in the class III BSC.*

## Introduction

Laboratory techniques require a combination of fine and gross motor skills, hand-to-eye coordination, and the ability to cope with imposed working restrictions. Of particular concern in the laboratory is the impact of working in personal protective equipment (PPE) on these skills. Despite meeting current European standards, many items of PPE may not be suitable for use in some working environments (Mayer & Korhonen, 1999). Use of PPE over long periods can restrict dexterity, reduce vision, and increase fatigue (Jones & Eagleson, 2001; Garner et al., 2004). In previous studies that assessed the risk of accidents, pegboard-style dexterity tests have quantified the levels of dexterity lost when using PPE and laboratory containment equipment (Sawyer et al., 2007; Sawyer & Bennett, 2005). The use of flexible film and half suit isolators were found to significantly reduce fine finger dexterity and gross and manual dexterity when compared to working in less restricted areas such as on a bench (Sawyer et al., 2007). Tasks requiring fine finger dexterity were found to be most affected by the use of barrier containment systems.

The pegboard dexterity tests, as originally developed

by employment agencies, were used in previous studies to select dextrous assembly workers. Dexterity is an ability that consists of several sub-factors (Fleishman & Hempel, 1954). It is possible that the type of dexterity required in assembly work may not translate to laboratory work. Although pegboard tests may indicate the effects of using PPE and containment equipment on dexterity, the best way to assess any effect on the accuracy of laboratory skills within containment would be to use tests that replicate the laboratory tasks conducted within containment equipment.

The two main types of BSCs used are open fronted (class I or II) cabinets and completely enclosed glove boxes (class III). The use of these cabinets and fume hoods has been associated with an increased risk of repetitive strain injuries (Hoskins & Erickson, 1998). Indeed Jones and Eagleson (2001) have designed a class II, Type A/B3 (now designated as type A2, per NSF/ANSI 49, 2004) BSC specifically to incorporate the ergonomic requirements of laboratory workers. During the design process, poor posture due to improper positioning of the chair and work surface, glare and eye strain due to incorrect lighting, and even environmental factors (lack of fresh air and inappropriate ambient temperature and humidity) have all been identified as risk factors that reduce productivity and increase worker injury.

This study aimed to develop a novel test that replicates the motor skills required within restricted laboratory environments (BSCs). This test was used to observe and quantify any difference in technical accuracy when working in class II and class III BSCs. The hypothesis was that work within the class III BSC would significantly impede motor skills and technical accuracy, compared to work within a class II BSC.

## Materials and Methods

### Participants

The participants for this study were all employees of the Health Protection Agency, Centre for Emergency Preparedness and Response (HPA CEPR). In this investigation into the effect of working in class II and III cabinets on technical accuracy, the 20 participants used a class III BSC regularly as part of their job. In order to determine any effects on pipetting accuracy, they were interviewed about the length of time they had worked in class III BSCs, and their responses were recorded. Ages of participants were also recorded.

## Biological Safety Cabinets: Design and Ergonomics

There are three main types of biological safety cabinets: class I, II and III.

**Class I and Class II Safety Cabinets**—Class I and II BSCs are of similar construction, but the air flow pattern within the BSC determines whether a cabinet of this type is classified as I or II, based on whether protection is required for the operator only, or for both the product and the operator. Horizontal movement in class I and II BSCs is unrestricted. Only vertical movement is restricted by the height of the opening. Work is carried out through a front opening. The hands and arms can be moved freely within the area and are restricted only by the screen at the front. In this study, a class II BSC (Walker Safety Cabinets) built to BS5726:2000) was used as a representative open fronted BSC (Figure 1).

**Class III Safety Cabinet**—Class III BSC is a closed-fronted, self-contained, negative pressure cabinet. Work inside the BSC is done through attached arm length latex (may also be neoprene, butyl rubber or Hypalon, consider a more generic term like “rubber”) gauntlets. The glove ports are in fixed positions that restrict horizontal and vertical movement and may be very uncomfortable. The class III BSC used in this study was built to a HPA CEPR design by Hitech (Salisbury, UK) and is shown in Figure 2.

The fine motor skills of an operator may be reduced by working in a class III cabinet when compared to a class I or II BSC. A novel test was required to compare the accuracy of work in class II and III cabinets.

### Pipetting Accuracy Test

**Test Administration**—The same administration technique was used during the validation of the test and the

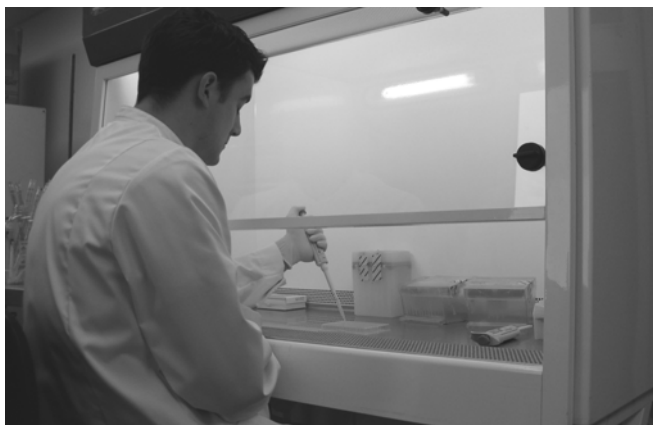
subsequent comparison of the safety cabinets. Participants were seated, at a bench, or at a BSC and wore latex laboratory gloves (Kimberly-Clark Safeskin satin plus SP2330E) of their preferred size. They used a single channel calibrated pipette (Labsystems Digital Finnpiquette 40-200 $\mu$ L) to complete three serial dilutions across a row of 12 wells in a flat bottomed 96 well plate (Scientific Laboratory Supplies, Sterilin microtitre plate, non-sterile, product number 79804). 100 $\mu$ L of sterile water (water for irrigation, Versol or sterile H<sub>2</sub>O) was pipetted into each of the 12 wells. 175 $\mu$ L of food dye (Supercook green food colouring) was added to the first well and mixed by alternately aspirating and expelling the contents two or three times using the pipette. Then, 175 $\mu$ L was taken from the first well and added to the next well and mixed. This serial dilution was continued along the row for all 12 wells. 175 $\mu$ L was taken out of the last well and discarded. The pipette tip was then replaced and the same procedure was repeated for the second and third row. The optical density of the plate was determined using an automated plate reader (Anthos reader 2001, Anthos Labtech Instruments, Life Sciences International) at a wavelength of 450nm, and the optical density recorded for each well.

Participants were instructed that they would be timed while completing the task, but to take as much time as necessary and focus only on accuracy, not on speed. The total time elapsed to complete the three rows of serial dilution was recorded, from the initial point of drawing up water into the pipette tip, to disposal of the last pipette tip.

**Instructions to Participants**—The instructions to participants were identical for both phases (validation of the pipetting test and the comparison of work within the class II and III BSCs). The participants’ instructions to complete the pipetting test were as follows:

**Figure 1**

A class II cabinet has a large front opening allowing a wide range of horizontal movement of the hands and arms of the operator.



**Figure 2**

A class III cabinet has gloves attached in a fixed position. This restricts the horizontal range of movement for the operator.



“Pipette 100 $\mu$ L of water into 12 wells across row A of a flat bottomed 96 well plate. Then pipette 175 $\mu$ L of neat food dye into the first well and mix two or three times with the pipette. Take 175 $\mu$ L out of this well, add to the next well and mix 2-3 times. Repeat across the row, until the end well. Take the last 175 $\mu$ L out and discard. Change pipette tips and start on the next row, first filling the row with 100 $\mu$ L of water, then adding and serially diluting the food dye.”

### Test Reliability and Validation

Participants in the reliability tests all had current laboratory experience unless otherwise stated.

**Plate Reader Reliability**—Plate reader reliability, always of concern, was assessed in this study by using the reverse plate wet test. Repeated readings of a dye loaded plate in normal and reversed positions allowed for analysis of the variation independent of pipetting error, or other experimental variables (Harrison & Hammock, 1988). A plate containing the same dilution of dye across all the wells was read forwards, then reversed and read again. Thus, each well had two readings that should have been identical.

**Test Validation**—Criterion-based validation requires a statistically significant relationship to be demonstrated between test performance and job performance. The simplest method for assessing the validity of this test was to compare the inaccuracy scores from the pipetting test between workers who used pipettes in laboratory work on a regular basis with those who had never used a pipette before. The hypothesis was that if the pipetting test was a true measure of pipetting accuracy, then experienced participants should demonstrate a significantly higher degree of accuracy than those who had never used a pipette before. The administration procedure was altered for the criterion-based validation test so only two rows of serial dilutions were used for the comparison of experienced and inexperienced participants. This alteration was introduced due to test time constraints on the inexperienced participants who generally took much longer to complete the test.

For this comparison, the inexperienced participants (n=10) were office staff members and non-laboratory manual workers, whereas the experienced (n=10) participants were those who completed serial dilutions, or worked with 96 well plates on a regular basis. The inexperienced participants were first shown the correct use of the pipette by explaining and demonstrating to them how to take up water using the pipette and then to expel it into a well. They were then asked to pipet water into two wells anywhere on a 96 well plate using this method. Following this training, they were taught the serial dilution task and asked to conduct a practice dilution for five wells as explained to them, and were corrected if they made errors. This training took no longer than 15 min-

utes and was to ensure the pipette was handled correctly and not damaged during the test.

**Test Reliability**—For validation, the pipetting test was completed on the bench, not in a BSC. Participants completed the test on the bench several times within one day and also on different days to determine intra- and inter-assay variation.

Ten participants were used to calculate the test-retest coefficient derived by testing the same group of subjects two times, a week apart, in order to counteract any learning effect. The two sets of results were then correlated.

### Testing Participants Using Biological Safety Cabinets with the Pipetting Test

**Testing Class II and III Cabinets**—Twenty participants completed the pipetting test in both class II and III BSCs. Half of the participants completed the test in the class II BSC first, and then in the class III BSC. The other half completed the test in the class III BSC first, followed by the class II BSC in order to counteract any order effects.

**Statistical Analysis**—The data were analyzed using a statistical software package (Minitab Release 13.32; Minitab Ltd). For all calculations, only the optical density readings from well columns 5-12 on the plate were used in the results as the readings from well columns 1-4 were too concentrated (out of range for the plate reader used).

To determine intra-assay variation, and within and between day inter-assay variation, the serial dilutions for each test were completed in triplicate. The coefficient of variation (CV) was calculated between the triplicates for each point on the standard curve. The mean of these CVs was then taken to determine the variation in pipetting inaccuracy for each test. The test-retest coefficient to determine test reliability was calculated using Pearson's correlation coefficient. A paired t test was used to compare the scores between those who were experienced in pipetting and those subjects who were not experienced in pipetting in order to assess criterion-related validity.

**Calculating Pipetting Inaccuracy Scores**—To determine whether a participant's results deviated from an accurate standard curve, the optical density readings for an accurately pipetted serial dilution of food colouring were required. A representative standard curve was determined from the mean of the results from three participants who could complete the serial dilution with a high degree of accuracy (defined as consistently producing a coefficient of variation less than 3% between triplicates when pipetting).

The mean of each triplicate from the participants' test results was calculated using the formula given below. For each point on the standard curve, the participants' score (mean of the triplicate results) was subtracted from the corresponding point on the representative standard curve. The value obtained was then made absolute (i.e., a negative number was made positive; a positive value was

left unchanged) as only the magnitude of the difference between the participant's curve and the representative standard curve was required for this study. The inaccuracy was measured in arbitrary units (AU). This sum was divided by the number of points across the standard curve, in this case 8, to obtain the pipetting inaccuracy score.

Pipetting inaccuracy =

$$\frac{|a_5-b_5| + |a_6-b_6| + |a_7-b_7| + |a_8-b_8| + |a_9-b_9| + |a_{10}-b_{10}| + |a_{11}-b_{11}| + |a_{12}-b_{12}|}{8}$$

- a = representative standard curve result
- b = participants' serial dilution result (mean of triplicate scores)
- 5 = column number across the 96 well plate (the first four columns across the plate are excluded because the optical densities are out of range of the plate reader used)

Data were checked for parametric assumptions before analysis and transformed where necessary.

## Results

### Test Reliability and Validity

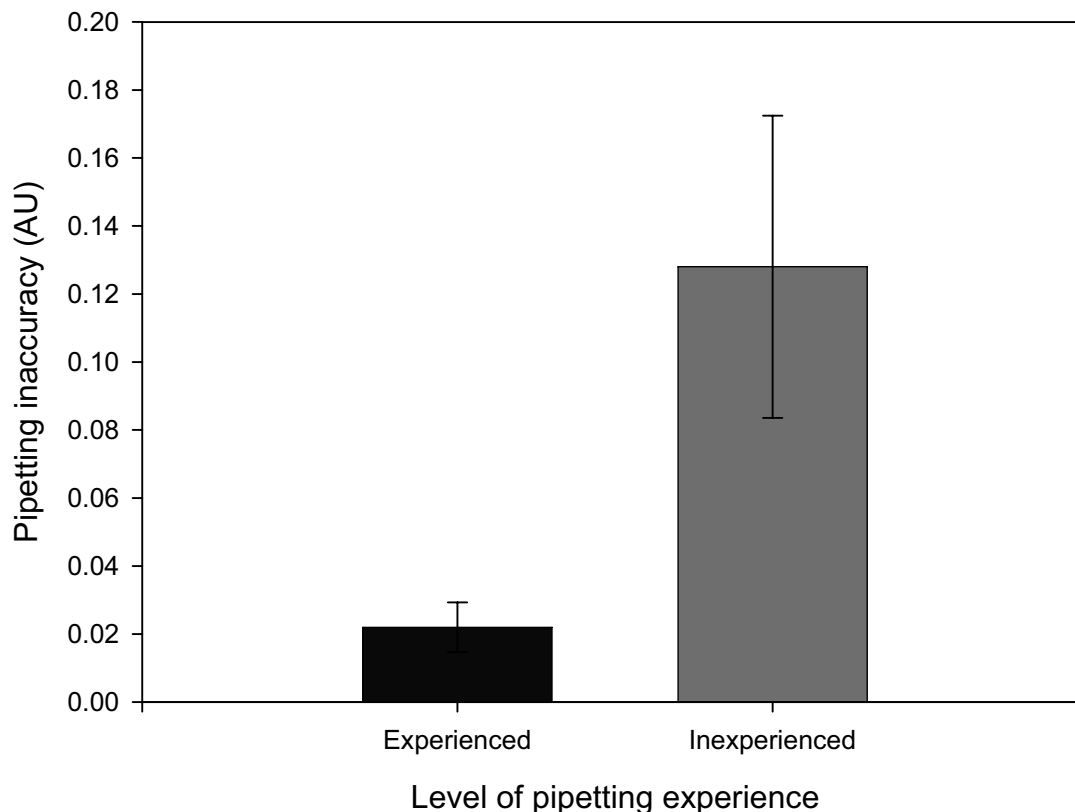
**Plate Reader Reliability: The Reverse Plate Wet Test**—A paired t test found no significant difference between the two sets of reverse plate wet test optical density readings.

**Test Validation**—To meet the parametric assumption of a normal distribution, the data were transformed ( $\lambda=0.5$ ) using a two sample t test. The experienced participants were found to perform significantly better on the pipetting test than the inexperienced participants ( $t=2.4$ ,  $p=0.028$ ). The difference between the two groups is illustrated in Figure 3.

**Test Reliability**—The pipetting test had an intra-plate variation of 1.4%, inter-plate (within day) variation of 3% and a between day variation of 3.6%. To determine the test-retest coefficient, 10 individuals were tested on one day and then tested again a week later. The Pearson's correlation coefficient between the two sets of scores was 0.868 ( $p=0.001$ ). Although participants were told to take as long as they needed during the test and to focus only on accuracy, the time taken to complete the test was recorded. The test-retest coefficient for time taken was also statistically significant (Pearson's correlation = 0.922,  $p<0.0001$ ).

**Figure 3**

The inaccuracy scores for experienced and inexperienced participants completing the pipetting test (working on a bench).



### Testing the Ergonomic Limitations of Class II and III Cabinets

**Participants**—The pipetting test was completed by 20 participants working in both class II and III BSCs. The age, gender, and distribution of experience are shown in Table 1.

**Pipetting Inaccuracy Scores and Times**—The participants had a range of inaccuracy scores, (0.009 AU to 0.273 AU). The time taken to complete the test had a range of 4 minutes 57 seconds to 18 minutes 49 seconds, with a mean time of 6 minutes and 54 seconds, and a standard deviation of 2 minutes and 55 seconds. Both the highest and lowest inaccuracy scores and completion

times were obtained from participants working in the class III BSC. The data were normalized using a  $\lambda$  value of -1. Using a two-sample t test, it was found that there was no significant difference between the class II and III BSC inaccuracy scores (Figure 4).

Ten experienced workers performed the test on a bench (working outside of a BSC) during the test validation. A comparison was then made between the pipetting inaccuracy scores from working outside a BSC (on a bench) to those scores from working in a class II and class III BSC. The data did not follow a normal distribution so were transformed using a lambda value of -0.5. Using a one-way analysis of variance (ANOVA), it was found that

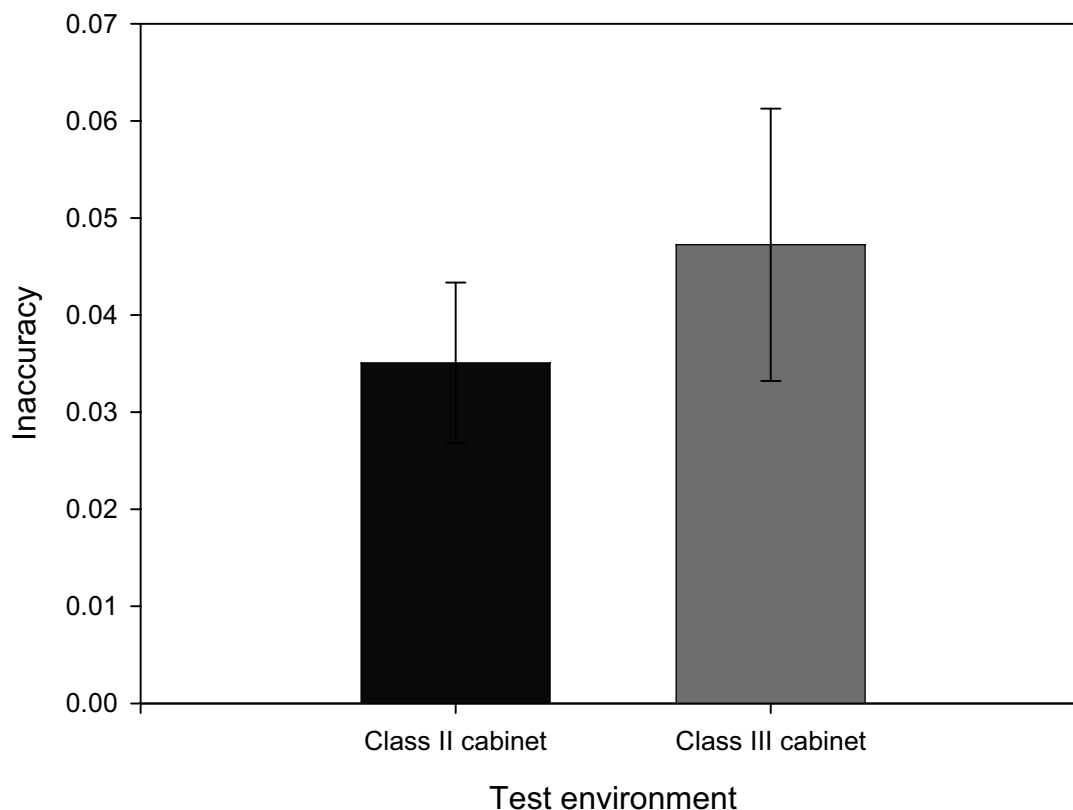
**Table 1**

The distribution of 20 participants used to test the difference between class II and III cabinets.

Factor	Distribution
Age	22-59 (mean = 32.6)
Gender	8 males; 12 females
Handedness	19 right handed; 1 left handed
Glove size	6 small; 10 medium; 4 large
Class III cabinet experience	1-126 months (mean = 60 months)

**Figure 4**

Inaccuracy scores for the pipetting test for class II and III cabinets.



there was a significant difference ( $f=3.94$ ,  $p=0.026$ ). Tukey's post hoc testing further showed that the difference was between working on a bench and in a class III cabinet ( $ci=0.248$ ,  $4.933$ ). There was no difference between working on a bench and working in a class II BSC. This finding concurred with previously reported results of dexterity testing (Sawyer et al., 2007).

The time taken to complete the pipetting test in the class II BSC was significantly shorter (Figure 5;  $t=2.96$ ,  $p=0.008$ ). There was a 9.8% increase in the time taken to complete the pipetting test in the class III BSC as compared to the class II BSC, which was a statistically significant, but small, difference.

**Class III Experience**—Using Pearson's correlation coefficient, there was no statistical correlation between experience and accuracy, or experience and time taken to complete the test in the test results from either cabinet type.

**Age**—There was no statistical correlation between age and accuracy, or age and time taken to complete the test in the test results from either type of BSC.

**Gender**—Using a two sample t test, there was no significant difference in inaccuracy, or time to complete the test between the genders for either type of BSC.

**Glove Size**—Using repeated measures one-way ANOVA to analyze inaccuracy by glove factor, no effect of glove size, on the inaccuracy scores measured by the pipetting test, was found. Using repeated measures one-

way ANOVA to analyze the time taken to complete the pipetting test by glove size, no effect of glove size was found on recorded times.

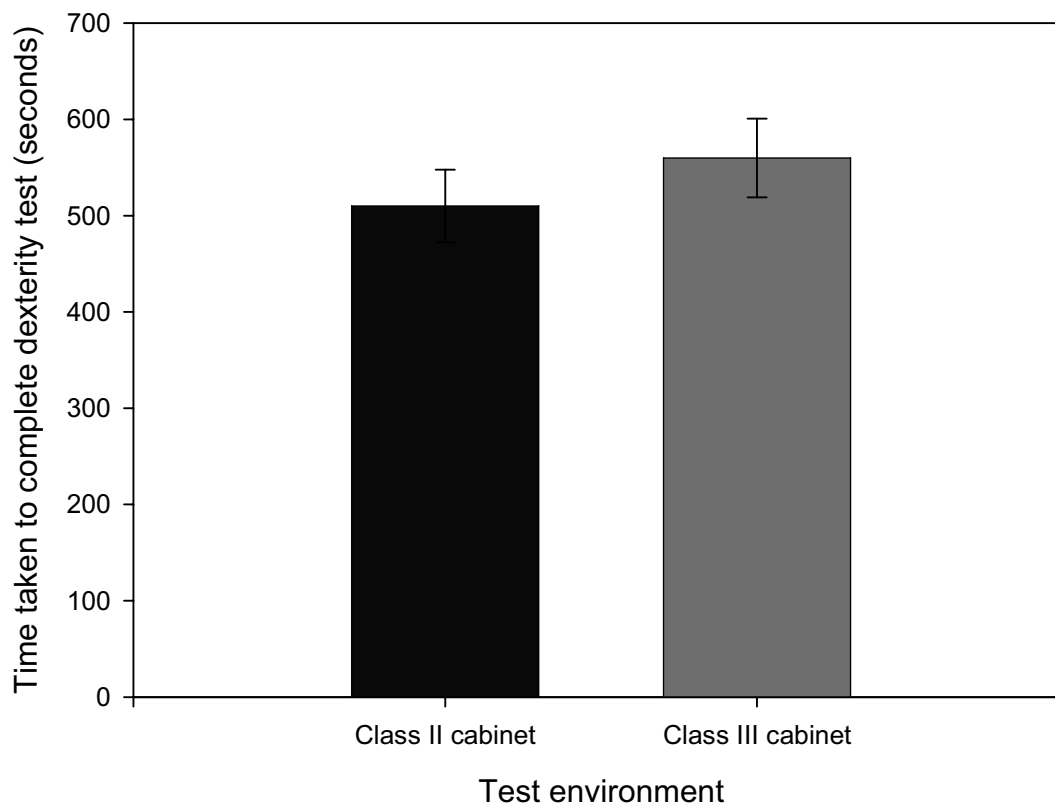
**Individual Scores**—There was a positive, but weak significant correlation between each participant's inaccuracy scores in the two types of BSCs (Pearson correlation = 0.636,  $p<0.003$ ). The times taken by each participant to complete the pipetting test also correlated between the two types of BSCs (Pearson correlation = 0.887,  $p<0.001$ ).

## Discussion

The pipetting test is a novel technique developed to assess the effect of containment equipment on the fine motor skills of the operator. The reverse plate wet test demonstrated the reliability of the plate reader used for the study. Laboratory workers who were experienced in using a pipette were found to have lower inaccuracy scores (mean 0.02) than the inexperienced office and manual workers who had never used a pipette. Those who had never used a pipette before had significantly higher inaccuracy scores (mean 0.13). These findings confirmed the hypothesis that the pipetting test was a valid measure of pipetting ability. The test was then used to assess pipetting abilities of participants working in the restricted environments of class II and III BSCs. The comparison found that the mean pipetting inaccuracy score

**Figure 5**

Average time taken to complete the pipetting test for the class II and III cabinets.



while working in the class III BSC was not significantly different from that obtained in the class II BSC, but was significantly greater than that determined when working on the bench (using a different test group). The mean time taken to complete the pipetting test was 9.8% faster in the class II BSC as compared to the class III BSC.

Contrary to the hypothesis, some of the participants were more accurate and indeed worked faster in the class III BSC than in the class II BSC. This finding may be due to familiarity of several participants with working in class III BSCs, lacking class II BSC experience. Two participants stated a preference for working in the class III BSC, because it allowed a more comfortable working position. The class III BSC user could rest his or her forearms on the glove ports whereas the class II BSC does not allow the worker to rest his or her arms because this action would obstruct the air grilles.

Participants were instructed to take as much time as they required to complete the test, and to work as accurately as possible while their test completion times were being recorded. The mean completion times for the class III BSC test were significantly longer than the mean class II BSC completion times. Although the difference was statistically significant, it was small, at 9.8%. It was concluded that participants heeded the class III BSC training given at the HPA where they are instructed not to rush class III BSC work, as their haste may reduce the accuracy of their work and increase the risk of accidents occurring.

A strong positive correlation for pipetting inaccuracy existed between the types of BSCs used by each participant and also for time taken. Thus if a participant were particularly slow, or inaccurate in a class II BSC, he or she would have been likely to have performed equally poorly in a class III BSC.

Based on age, gender and handedness, the descriptive statistics indicated that participants in this study represented a cross section of HPA CEPR staff that used class III cabinets. All had class III BSC work experience of at least one month's regular use. Only one participant had no previous experience with the type of pipette used in the test. Several participants had never worked with multi-well plates before. However; in an unexpected finding, accuracy and time did not correlate significantly with the participants' level of class III BSC work experience. Participants were asked how many months, or years of experience they had and this value was recorded as an absolute figure. Some of the participants may have worked in a class III BSC continuously whereas others, with a longer history of BSC experience, may have used BSCs intermittently.

Age and gender also did not influence inaccuracy and time taken to complete the test. From previous studies (Sawyer et al., 2007; Sawyer & Bennett, 2005), it was thought that glove size would affect the results of this test. Gloves, which are too big for the wearer, may impede dexterity. In the class III BSC, large (size 9) gauntlets were

used for every participant. Participants wore latex, or nitrile laboratory gloves of their chosen glove size under these gauntlets. Thus, participants who selected large laboratory gloves should have had an advantage over those who wore smaller gloves. Previous studies with gloves have shown that overly large gloves impede fine finger dexterity. Fine finger movements are not needed when carrying out a serial dilution using a pipette. That may explain why the gloves size had no effect on time, or accuracy in this study.

If sufficient time is given to complete work, working in the class III BSC will have no more effect on an operator's technical accuracy than working in a class II BSC. The wide range of pipetting inaccuracy scores and of times taken to complete the test have shown that the type of BSC used has no real effect on technical accuracy, but there is a wide variation in technical ability between experienced operators.

## Conclusion

The pipetting test was developed to measure the accuracy of laboratory skills and was validated successfully. The test showed no difference in the levels of technical accuracy of laboratory work conducted in class II and III BSCs. However, operators worked slightly more slowly in a class III BSC when compared to work in a class II BSC. Levels of laboratory technical accuracy only differed between individual workers. No external factors such as age, glove size, experience, or gender, were identified that significantly affected all users to the same extent.

## References

- Fleishman, E., & Hempel, W. E. (1954). A factor analysis of dexterity tests. *Personnel Psychol.*, 7, 15-32.
- Garner, A., Laurence, H., & Lee, A. (2004). Practicality of performing medical procedures in chemical protective ensembles. *Emerg. Med. Australas.*, 16(2), 108-113.
- Harrison, R., & Hammock, B. D. (1988). Location dependent biases in automatic 96-well microplate readers. *J. Assoc. Off. Anal. Chem.*, 71(5), 981-987.
- Hoskins, D., & Erickson, J. (1998). Laboratory ergonomics. The wake up call: a case study. *Chem. Health Safety, Jan/Feb*, 22.
- Jones, R. J., & Eagleson, D. (2001). Ergonomic considerations in the development of a class II, type A/B3 biological safety cabinet. *Am. Clin. Lab.*, 20(4), 37-42.
- Mayer, A., & Korhonen, E. (1999). Assessment of the protection efficiency and comfort of personal protective equipment in real conditions of use. *Int. J. Occup. Saf. Ergon.*, 5(3), 347-360.
- Sawyer, J., Bennett, A., Haines, V., Elton, E., Crago, K., & Speight, S. (2007). The effect of microbiological containment systems on dexterity. *J. Occup. Environ. Hyg.*, 4 166-173.
- Sawyer, J., & Bennett, A. (2005). Comparing the level of dexterity offered by latex and nitrile safeskin gloves. *Ann. Occup. Hyg.*, 50(3), 289-296.