The Hawthorne effect on adherence to hand hygiene in patient care: a systematic review

Edward Purssell¹
Nicholas Drey¹
Jane Chudleigh¹
Sile Creedon²
Dinah J Gould¹

¹ School of Health Sciences, City, University of London, London UK
² School of Nursing and Midwifery, University College Cork, Cork, Republic of Ireland

Corresponding author: Edward Purssell
School of Health Sciences, City, University of London, 1 Myddleton Street, London EC1R 1UW.
Telephone 07782 374217
Email edward.purssell@city.ac.uk

Running title: Hawthorne Effect & Hand Hygiene
The Hawthorne effect on adherence to hand hygiene in patient care: a systematic review

SUMMARY
Background Numerous studies demonstrate that the Hawthorne effect (behaviour change caused by awareness of being observed) increases health workers’ hand hygiene adherence but it is not clear if they are methodologically robust, magnitude of the effect, how long it persists or whether it is the same across clinical settings.

Objective Determine rigour of the methods used to assess the Hawthorne effect on hand hygiene, effect size estimation, variations between clinical settings and persistence.

Methods Systematic literature review with meta-analysis.

Results Nine studies met the criteria for the review. Methodological quality was poor. Data pooling was possible across six studies. The Hawthorne effect ranged from 4.2% to 65.3% with a median of 35.6%. It was 4.2% in one study conducted in intensive care and 16.4% in transplant units. It was most marked when data were collected across an entire hospital and in a group of general hospitals. Differences between wards in the same hospital were apparent. In the two studies where duration was estimated, the Hawthorne effect appeared transient.

Conclusions Despite methodological shortcomings the review indicates clear evidence of a Hawthorne effect on general wards. There is some evidence that it may vary according to clinical specialty and across different wards within the same organisation. The review identifies a need for standardised methodologies to measure the Hawthorne effect in hand hygiene to overcome the dilemma of reporting the potentially inflated rates of adherence obtained through overt audit. Occasional covert audit could give a better estimation of ‘real’ hand hygiene adherence but its acceptability and feasibility to health workers need to be explored.

Words in abstract = 236
INTRODUCTION

Hand hygiene is essential to avoid transmission of nosocomial pathogens [1] and helps prevent spread of community-acquired infection in settings where health care is delivered.[2] In many countries health workers’ hand hygiene is audited routinely, usually by direct observation and manual documentation and is regarded as a key marker of the quality of care. Overt manual audit increases adherence to hand hygiene protocols and is most marked when auditors are known to health workers.[3] This is a manifestation of the Hawthorne effect: increased productivity when individuals are aware of scrutiny, either in the workplace or when they take part in research.[4] Other inaccuracies encountered during overt audit include data loss through poor vantage, bedside curtains obscuring clinical activity and failure to document all hand hygiene opportunities and events.[5] The Hawthorne effect is a major source of bias when overt audit takes place [6] and is the single greatest methodological hurdle reported by research teams attempting to evaluate the effectiveness of interventions to promote hand hygiene adherence.[7] Managers and health workers know about the Hawthorne effect and dismiss overt audit as a valid reflection of practice.[8] Hand hygiene data that do not appear credible to health workers are unlikely to change behaviour in relation to adherence.[9]

A number of methods have been adopted to overcome the limitations of overt audit but all have drawbacks. Product consumption is not a valid measure of hand hygiene behaviour [5] while covert manual audit is subject to the same challenges of data loss as overt manual audit and is not recommended by the World Health Organization (WHO) [10] because it can promote mistrust and resentment if health workers become aware that it is taking place and the practice could be viewed as ethically unsound. Different types of electronic hand hygiene monitoring systems (EMSs) are available.[11] Most track adherence only in relation to Moments 1, 4 and 5 [12] of the WHO’s Five Moments for Hand Hygiene [13] and their ability to detect hand hygiene opportunities and events can be affected by the health worker’s location and positioning of body-worn electronic sensors.[14] Accuracy can be greater in simulated settings than hospital wards [14] and system malfunction can result in data loss. Habituation might reduce the behavioural impact of being observed [15] but is rarely employed in hand hygiene research.[16] Overt audit is considered to be the ‘gold standard’ because it allows the hand hygiene event to be evaluated in the
context of patient care and provides an opportunity for correction and improving practice not possible when other methods are employed.[10] Direct observation, whether overt or covert, enables the auditor to assess the quality of hand hygiene technique but not when an EMS is employed.[5] Thorough hand hygiene events allowing adequate contact of all hand surfaces with the antiseptic agent is essential to remove nosocomial pathogens.[17]

There is urgent need to establish magnitude of the Hawthorne effect in hand hygiene to interpret the findings of overt manual hand hygiene audit and inform practice and policy. The aims of this systematic literature review were to:

1. Determine the rigour of the methods used to determine the Hawthorne effect.
2. Estimate size of the Hawthorne effect in patient care settings and identify any variations between patient settings.
3. Determine how long the Hawthorne effect persists.
4. Evaluate the effectiveness of any interventions to minimise the Hawthorne effect.
5. Identify the cost of interventions used to minimise the Hawthorne effect.

METHODS

Search strategy and study selection

Medline and Embase were searched with the terms: ‘Hawthorne effect’ and ‘Hawthorne effect’ AND ‘hand’. We also identified potentially eligible papers from personal collections held by members of the research team. We established the most high-yield journals publishing relevant papers and hand-searched these. Reference lists of all retrieved studies were hand-searched. To be eligible for inclusion papers had to report the results of hand hygiene monitoring where data were collected by routine overt audit and a comparator (e.g. EMS, covert manual observation, closed circuit television) at the same time. Our objective was to identify any differences between what usually takes place during routine clinical practice and the comparator, not to explore differences between a comparator and overt observation introduced especially for the study that might not reflect real life. Synchronous data collection was essential to ensure that the datasets were directly comparable. Hand hygiene opportunities and adherence are influenced by clinical workload, the nature of the activity undertaken and interruptions. This meant that when the comparator was an
EMS, the study could only be included if the automated data obtained during periods when manual overt audit was not in progress were excluded from analysis. We included only those studies where the same criteria used to identify hand hygiene opportunities and adherences were applied in both audit methods. Where an EMS was employed we obtained details of data capture from website information or from manufacturers if it was not reported in the publication. We also attempted to establish whether the algorithms used to identify hand hygiene opportunities and adherence by EMSs were reported and how they had been agreed when the system was designed. Studies reporting product consumption were excluded unless an additional comparator was employed.

Data extraction and synthesis
Publications meeting the above criteria were read in depth to assess the rigour with which the comparator had been validated. Indicators of good practice during covert manual audit would be employing data collectors unknown to staff, health workers remaining unaware of scrutiny, minimal data loss, training auditors, reporting acceptable (>80%) agreement at inter-rater reliability testing and periodically assessing and revalidating auditors. We planned to apply the same criteria to data analysis in studies where CCTV footage was inspected. For studies taking electronically obtained data as the comparator we planned to assess whether periods of data loss were acknowledged and excluded from analysis and how validity of the EMS had been determined. The data extraction proforma was developed by DJG, EP and SC. Two members of the research team worked together to select the included publications and assess quality (DJG and EP). ND checked eligibility of the studies and data extraction but third party arbitration to resolve divergent opinion was not required. The forest plot was produced using the R package meta.[18]

RESULTS
Literature review
Forty eight potentially eligible full text publications were identified through electronic searching, six were held in personal collections and two were identified by hand-searching. We excluded conference abstracts because they contained too little detail to withstand critical appraisal. Of the potentially eligible studies, 18 did not explore the Hawthorne effect in relation to hand hygiene in patient care or did not contain
empirical data. Twelve further studies were excluded because the data were not collected synchronously.[19 - 30] Four studies were excluded because they compared two covert audit methods.[31 – 34] Three studies were excluded because overt audit and the EMS did not collect the same data although synchronous monitoring occurred.[35 – 37] One study was excluded because it was unclear whether the comparator was overt or covert [38] and another publication was excluded because overt manual audit was introduced especially for the study.[39] This is shown in Figure 1.

[INSERT FIGURE 1 HERE]

Figure 1. PRISMA Flow-chart

Nine studies reported the results of hand hygiene audit with synchronous data collection conducted by routine overt audit and at least one comparator.[40 -48] In two studies the comparator was covert manual audit.[42 45] In five studies the comparator was an EMS.[40 - 42, 45, 48] In one study there was comparison of overt manual audit with covert manual audit and an EMS [46] and in one study CTTV was used.[43] Data in the included studies were collected across an entire hospital by covert manual observation.[42] five general hospitals where the comparator was data collection by covert manual audit, [45] medical and surgical wards in a tertiary hospital where the comparator was data collection by an EMS [44] and an adult step-down unit where the comparator was data collection by an EMS.[40] Differences between medical and surgical wards in the same hospital were apparent in one study. [44] Table 1 shows that for the remaining studies details of validation were unclear or not comprehensive. The included studies were undertaken in Australia, Brazil, Canada, the Middle East and the United States (see Table 1). One study was reported from northern Europe and none from the United Kingdom. All took place in acute care settings. In one study [45] data collection involved five hospitals. The remaining studies involved a single organisation with data collection restricted to one ward or a small number of wards, often of a highly specialist nature. In some studies the number of hand hygiene episodes documented was considerable while in others it was comparatively small, ranging from 911,791 [47] to as few as 659.[40] None of the included studies presented a comprehensive account of the steps taken to validate the
In three studies [43, 44, 47] the authors reported that this information had been presented in an earlier publication and cited it in the text. In two of these studies the earlier publication related specifically to the EMS in question. [28, 49] In the other case the cited publication contained non-specific information relating to which of the Five Moments EMSs are generally able to identify. [9]

**Meta-analysis**

It was possible to extract or calculate the effect sizes for the difference between overt and comparator estimates of hand hygiene in six of the included studies. [40, 42 - 45, 47] In the other studies estimates of the Hawthorne effect were not reported or could not be calculated. In four studies it was possible to calculate the Hawthorne effect with confidence intervals, these are shown in Figure 2. In addition to these, two studies published only an estimate of the Hawthorne effect. One study established a difference of 37.31% covert manual and overt audit in the hour before overt observation and 53.33% in the hour afterwards. [40] The other study [45] established a weighted difference of 29.7% between overt audit and the comparator and 33.8% after feedback.

[INSERT FIGURE 2 HERE]

**Figure 2. Estimates of the Hawthorne Effect**

Two studies investigated duration of the Hawthorne effect. [40, 47] In one study hand hygiene decreased by 53.3% in the hour after overt observation ceased. [40] In the other study [47] hand hygiene adherence rate was 4.08 per hour pre-audit, rising to 5.72 during the overt audit period. Less than an hour afterwards adherence fell to 5.6 hand hygiene events per hour. One hour later the rate was 4.06 events per hour falling to 3.9 events per hour after two hours.

Using the Newcastle Ottawa Scale [50] the risk of bias within and across the studies was generally low, although these studies did not fit conventional definitions of intervention or cohort studies. The main risk is that associated with studies being conducted primarily in single units or hospitals. Due to the nature of the research question the same cohort were studied, and the aim of the studies was to compare the
two methods of assessment and to quantify the bias from observation compared to automated methods. The full assessment is shown in Supplemental File 1.

DISCUSSION

Ours appears to be the first systematic literature review to explore the impact of the Hawthorne effect on hand hygiene adherence in patient care. Of 30 potentially eligible studies only nine were sufficiently robust to meet the inclusion criteria. Of these it was possible to extract comparable data from six studies. The included studies show wide variations in the estimate of the Hawthorne effect possibly for methodological reasons or because they reflect different attitudes towards hand hygiene between clinical areas. In the study [43] reporting a negative effect with hand hygiene lower when measured by the comparator, the findings were based on small samples of hand hygiene episodes. There were differences between medical and surgical wards in the same hospital in the only study where it was possible to explore such variation.[44] Duration appeared short-lived in the two studies exploring persistence of the Hawthorne effect.[40, 47] No research team explored economic considerations although those using an EMS as the comparator pointed out that automated systems are expensive. Quality of the hand hygiene event was not assessed in any of the studies although it would have been possible in those where covert manual audit was employed.

Limitations exist in relation to our review and the empirical studies we included. ‘Hawthorne effect’ is not a MeSH search term in Medline. It is the entry term for the much broader Effect Modifier Epidemiologic MeSH term. It is therefore possible that some studies were not traced but unlikely as members of the research team were very familiar with the wider literature on hand hygiene adherence. Of the papers held in our personal collections not appearing in the searches, only one met the inclusion criteria. Overall the empirical studies were poorly reported, the number of hand hygiene events observed varied, most studies included only a few wards in the same organisation and lacked external validity. The lack of a specific MeSH term may be problematic for future research teams addressing the Hawthorne effect.
Originally we planned to restrict the review to studies where a robust comparator had been applied. This proved impossible because reporting was poor and the studies suffered from failure to prepare covert auditors adequately or ensure the validity of the EMSs. For example, Kovacs-Litman [42] reported only that covert auditors were trained and collected data for ‘a short period’ in each clinical setting to avoid recognition, while in the study reported by Scherer [45] covert auditors received only one day of training and one day of supervised practice. Inter-rater reliability testing and revalidation of auditors were not mentioned in either study. Where EMSs were employed, details about the system were often scant. In two studies there was evidence of data loss [40, 45] and one research team admitted that covert data collectors might have been recognised.[45] The algorithm was disclosed in only one study where an EMS was used.[40]

Despite the above limitations, the review demonstrates clear evidence of a Hawthorne effect on general wards. It was lowest in studies reported from intensive care [43] and transplant units,[47] possibly because health workers in specialist, high risk settings are more aware of the need for hand hygiene and are more adherent. Marked differences before and after the introduction of an intervention to improve hand hygiene were also reported in an earlier study where data were collected in ITUs and acute care of the elderly wards.[51] In this study greater awareness of the need for hand hygiene and the additional training received by nurses in ITUs were thought to explain the difference. Hand hygiene adherence varies between wards and is influenced by local culture and leadership.[7] It is also possible that managers directed research teams to venues where practice was already good and it was anticipated that health workers would be more likely to tolerate unannounced covert observation or the introduction of an EMS. Differences between medical and surgical wards in the same hospital [44] show that within the same organisation, the nature of the clinical setting could be important.

Our review is timely. The Hawthorne effect is the major source of bias when overt audit of hand hygiene takes place [6] and inability to control for it was identified as the single most important methodological challenge confronting research teams evaluating the effectiveness of interventions intended to promote hand hygiene adherence.[7] Our review confirms that on general wards the Hawthorne effect
operates as a major obstacle when hand hygiene adherence is measured and that pooling data for the entire organisation may be misleading as there is evidence of considerable variation between wards in the same hospital. The review identifies a need for standardised methodologies to measure the Hawthorne effect. As anticipated we were unable to locate any studies where habituation was taken as the comparator.[16] In the other studies the comparator was either covert manual audit, an EMS or CCTV but there is insufficient evidence to recommend any as the best approach to assess the Hawthorne effect. At present it is possible only to recommend that irrespective of the comparator employed, detailed evidence of how it has been validated should be given.

Hand hygiene is audited routinely by overt methods in many countries. The inflated rates of adherence reported as a result of this resource-intensive exercise are often displayed on health providers’ websites providing false reassurance concerning an important patient safety issue. Health workers are aware that the high levels of adherence commonly reported lack validity [8] yet are obliged to discuss them with their staff. If adherence declines, reminders and in some organisations punitive action follow.[52] The findings of our review highlight the ethical issues surrounding the collection and use of data obtained with flawed methodologies.

Overall the quality of the studies available for review was disappointing. Over half the potentially eligible studies had to be discarded because of avoidable errors. These included failures to ensure that comparison was made between synchronously obtained datasets, failure to ensure that the same criteria were used to identify hand hygiene opportunities and adherences, omissions of reporting related to the validation of covert auditors and omission of key information concerning EMSs. Care should be taken to avoid the same mistakes in future studies. Further research is required to explore whether the Hawthorne effect varies between clinical settings and to establish whether there are patterns between the same types of clinical settings in different healthcare provider organisations. The relative merits of different approaches to audit systems acting as the comparator needs to be explored further to establish a standardised methodology to assess and allow for the Hawthorne effect in order to inform policy, practice and improve the rigour of interventions intended to improve hand hygiene. To provide maximum information these comparators should provide data in relation to the quality of hand hygiene and adherence in relation to Five Moments. The costs of undertaking routine overt hand hygiene audit do not appear to
have been estimated. Economic analysis is necessary to establish these costs and the costs of periodically employing a comparator to check the validity of overt audit data. Finally, as the Hawthorne effect is such a key research topic in relation to hand hygiene adherence and other epidemiological phenomena [4] it would be helpful if it could be given its own MeSH term separate from other confounding factors.

Better-controlled studies to assess magnitude and persistence of the Hawthorne effect applied to hand hygiene are required but as we explain above, these will be challenging to undertake and in the meantime a practical solution needs to be found to overcome this pressing clinical problem. It is also important to be aware of possible differences between clinical specialities, this may be particularly significant for those who use these data for clinical and policy decision making purposes. Covert audit is the most obvious solution. Most healthcare providers do not use EMSs, so occasional covert audit appears to be the most practical solution. This approach is not recommended in current policy (10) but there was no evidence of health workers complaining about covert audit in the studies we reviewed and any suggestions that it is unethical need to be balanced against the use of audit findings that are known to be invalid. Furthermore, these limitations need to be balanced against the need to protect patients and colleagues, as there is clear evidence from these data that their might be overconfidence in the level of current compliance.

Before such a major policy change is contemplated it will be necessary to obtain the views of clinical leaders to establish the acceptability and feasibility of occasional covert hand hygiene audit through focus groups or qualitative interviews. If the findings are positive, fieldwork can then be undertaken to establish practicalities: ensuring that health workers remain unaware that covert audit is in progress and that auditors are fully trained with good inter-rater reliability. If rigorously undertaken, covert audit will provide an estimation of the ‘real’ rate of hand hygiene adherence but care must be taken to avoid direct comparison with the outcomes of overt audit unless both audits are conducted at the same time.

In conclusion our systematic review demonstrates clear evidence of a Hawthorne effect for hand hygiene on general wards and shows that the findings of overt routine
hand hygiene audit are inflated. Better-controlled studies to assess magnitude and persistence of the Hawthorne effect applied to hand hygiene are required. …

Acknowledgments: None

Competing Interests: DJG has received funding from Essity relating to a research project. None of the other authors have any competing interests to declare.

Funding: No specific funding was received for this study.

Contribution statement: the study was conceived by DJG. DJG and EP undertook data extraction. All authors contributed to the analysis. DJG and EP drafted the manuscript to which all authors contributed.
REFERENCES


[17] Beggs CB, Shepherd SJ, Kerr KG. Increasing frequency of hand washing by healthcare workers does not lead to commensurate reductions in staphylococcal infection in a hospital ward. *BMC Infect Dis* 2008; **114**.


[23] Eckmanns, T., Bessert, J., Behnke, M., Gastmeier, P., & Ruden, H. Compliance with antiseptic hand rub use in intensive care units: the Hawthorne effect. *Infect Control Hospital Epidemiol* 2006; **27**: 931–934. [https://doi.org/10.1086/507294](https://doi.org/10.1086/507294)


Figure 1. PRISMA flow-chart
<table>
<thead>
<tr>
<th>Study</th>
<th>Observed Events</th>
<th>Electronic Events</th>
<th>Total Risk Difference</th>
<th>RD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>McLaws et al (2015) Medical 2015 Audit 2</td>
<td>155</td>
<td>52137</td>
<td>168416</td>
<td></td>
<td>0.65 [0.62, 0.69]</td>
</tr>
<tr>
<td>McLaws et al (2015) Medical 2015 Audit 3</td>
<td>189</td>
<td>54075</td>
<td>187234</td>
<td></td>
<td>0.65 [0.62, 0.68]</td>
</tr>
<tr>
<td>McLaws et al (2015) Medical 2015 Audit 1</td>
<td>177</td>
<td>76064</td>
<td>256728</td>
<td></td>
<td>0.63 [0.59, 0.66]</td>
</tr>
<tr>
<td>McLaws et al (2015) Medical 2014 Audit 2</td>
<td>141</td>
<td>49466</td>
<td>161728</td>
<td></td>
<td>0.57 [0.52, 0.62]</td>
</tr>
<tr>
<td>McLaws et al (2015) Medical 2014 Audit 1</td>
<td>333</td>
<td>49690</td>
<td>161728</td>
<td></td>
<td>0.54 [0.51, 0.58]</td>
</tr>
<tr>
<td>McLaws et al (2015) Medical 2014 Audit 3</td>
<td>194</td>
<td>40057</td>
<td>150024</td>
<td></td>
<td>0.53 [0.49, 0.58]</td>
</tr>
<tr>
<td>Kovacs-Litman (2016) Nurses</td>
<td>756</td>
<td>879</td>
<td>330</td>
<td></td>
<td>0.41 [0.37, 0.45]</td>
</tr>
<tr>
<td>McLaws et al (2015) Surgical 2015 Audit 3</td>
<td>268</td>
<td>27788</td>
<td>59110</td>
<td></td>
<td>0.37 [0.33, 0.41]</td>
</tr>
<tr>
<td>McLaws et al (2015) Surgical 2014 Audit 1</td>
<td>354</td>
<td>52015</td>
<td>99450</td>
<td></td>
<td>0.36 [0.32, 0.39]</td>
</tr>
<tr>
<td>McLaws et al (2015) Surgical 2014 Audit 2</td>
<td>228</td>
<td>65216</td>
<td>125415</td>
<td></td>
<td>0.34 [0.30, 0.38]</td>
</tr>
<tr>
<td>Kovacs-Litman (2016) Total HH performances</td>
<td>2769</td>
<td>3309</td>
<td>799</td>
<td></td>
<td>0.34 [0.31, 0.36]</td>
</tr>
<tr>
<td>McLaws et al (2015) Surgical 2015 Audit 2</td>
<td>252</td>
<td>76655</td>
<td>138050</td>
<td></td>
<td>0.31 [0.27, 0.35]</td>
</tr>
<tr>
<td>McLaws et al (2015) Surgical 2015 Audit 1</td>
<td>84</td>
<td>70864</td>
<td>130405</td>
<td></td>
<td>0.26 [0.20, 0.33]</td>
</tr>
<tr>
<td>McLaws et al (2015) Surgical 2014 Audit 3</td>
<td>285</td>
<td>72480</td>
<td>129305</td>
<td></td>
<td>0.25 [0.21, 0.29]</td>
</tr>
<tr>
<td>Kovacs-Litman (2016) Physicians</td>
<td>175</td>
<td>230</td>
<td>469</td>
<td></td>
<td>0.19 [0.12, 0.26]</td>
</tr>
<tr>
<td>Vaisman (2020) Overall</td>
<td>4034</td>
<td>7107</td>
<td>361670</td>
<td></td>
<td>0.16 [0.15, 0.18]</td>
</tr>
<tr>
<td>Lifshitz-Rivin (2019) After suspected exposure to fluids</td>
<td>11</td>
<td>29</td>
<td>12</td>
<td></td>
<td>0.13 [-0.08, 0.33]</td>
</tr>
<tr>
<td>Lifshitz-Rivin (2019) After patient contact</td>
<td>45</td>
<td>86</td>
<td>31</td>
<td></td>
<td>0.10 [-0.05, 0.25]</td>
</tr>
<tr>
<td>Lifshitz-Rivin (2019) Total HH performances</td>
<td>111</td>
<td>332</td>
<td>96</td>
<td></td>
<td>0.04 [-0.03, 0.11]</td>
</tr>
<tr>
<td>Lifshitz-Rivin (2019) Before patient contact</td>
<td>6</td>
<td>54</td>
<td>5</td>
<td></td>
<td>0.03 [-0.07, 0.12]</td>
</tr>
<tr>
<td>Lifshitz-Rivin (2019) After contact with patient surroundings</td>
<td>46</td>
<td>138</td>
<td>45</td>
<td></td>
<td>-0.04 [-0.15, 0.08]</td>
</tr>
<tr>
<td>Lifshitz-Rivin (2019) Before aseptic contact</td>
<td>3</td>
<td>13</td>
<td>10</td>
<td></td>
<td>-0.07 [-0.43, 0.30]</td>
</tr>
</tbody>
</table>

Hawthorne effect