Fire and rescue operational effectiveness: the effect of alternative crewing patterns

Maher, K., Bateman, N. & Randall, R.

Author post-print (accepted) deposited by Coventry University’s Repository

Original citation & hyperlink:
https://dx.doi.org/10.1080/09537287.2019.1701232

DOI 10.1080/09537287.2019.1701232
ISSN 0953-7287
ESSN 1366-5871

Publisher: Taylor and Francis

This is an Accepted Manuscript of an article published by Taylor & Francis in Production Planning & Control on 13/12/2019, available online: http://www.tandfonline.com/10.1080/09537287.2019.170123

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.
Fire and Rescue operational effectiveness: the effect of alternative crewing patterns

Abstract
Decreasing demand and reduced budgets are driving changes to traditional crewing methods in the UK Fire and Rescue Service. Using an integration of two years’ operational data within a novel framework, this paper assesses the impact of alternative crewing on operational effectiveness in one UK FRS. Changes in crewing patterns were implemented without substantial impact on overall operational effectiveness, but there may be a risk to wider operational resilience. The Overall Effectiveness of Fire Operations (OEFO) assessment tool can inform FRS decision making in an authentic way allowing stakeholder confidence in the outcomes, whilst being timely and not too complex or costly to evaluate. The OEFO approach is an important contribution to practice through its ability to assess public services at a time of challenging reform and demonstrates alterations can be made to crewing patterns to better match demand provided there is consideration of the potential wider impact.

Keywords
Overall Equipment Effectiveness, public sector performance measures, fire service, emergency services, operational effectiveness

Introduction
Operational effectiveness (OE) is defined as activities that enable an organisation to a) better utilise its resources, b) better implement its processes and c) achieve its mission and objectives (Porter 1996), and is a central feature of the continuous improvement of functional performance. A ‘performance revolution’ (Neely 1999) has transformed performance measurement for public services over the last thirty years (OECD 2004) and Bititci, Cavalieri, and Cieminski (2005) call for an update in how public services are measured. In the UK Fire and Rescue Service (FRS) annual key performance indicators, governed by the Department of Communities and Local Government (DCLG), exemplify the new public
management (NPM) ethos (Carvalho et al. 2006). NPM focuses on efficient use of resources, performance monitoring and greater internal and external accountability to address the requirements of a wide range of stakeholders (Kloot 2009). The revised National Framework for Fire and Rescue Services (Home Office 2018) outlines these requirements indicating the interests of efficiency and effectiveness for delivering public safety, and the need for an independent inspectorate to ensure accountancy and transparency. In 2016 Theresa May, the then UK Home Secretary, highlighted the continuing difficulties in governance and scrutiny of the FRS due to the lack of clear auditing practices and the limited available data between services and overtime. In parallel to the 2012 reforms carried out within the Police Force, May pushed for transparent and publicly available data to allow for comparisons across FRS (May 2016). The provision of such data may also provide the ability to assess operational performance within services overtime, which is of benefit when assessing the impact of organisational change processes.

In this paper we examine the measurement of operational activities in the context of a change to crewing arrangements in four fire stations at one UK FRS (The Counties FRS¹). We offer a model of operational effectiveness for the fire and rescue service that aims to meet the OECD (2004) criteria of being reliable for stakeholders, timely, and not too complex or costly to evaluate. The purpose of the paper is to explore the utility of the operational effectiveness model for measuring performance within the fire and rescue service, and provide commentary on the impact of the new crewing arrangements on operational service delivery.

¹ The host organisation has been provided with the pseudonym ‘The Counties FRS’ to protect anonymity
Operational effectiveness in the Fire and Rescue Service

Little rigorous research has been conducted on operational effectiveness within the UK FRS, with only a small number of studies focusing on response times (Taylor 2016; Bateman, Maher, and Randall 2016; Jaldell 2005). The overall objective for the fire and rescue service is the ability to protect the public from harm and respond to emergency incidents, which involves complexity beyond just response times. Response times were used to assess the impact of station closures within London Fire Brigade, following Central Government’s cuts to public services (Taylor 2016), but poor variability in response times limits the use of this variable as a measure of effectiveness (Carvalho et al. 2006; Jaldell 2005). Previous work has attempted to quantify FRS performance in terms of input and output to fill the gap in efficiency and productivity measurement within this sector (Jaldell 2005). Confounding variables such as road and traffic condition limit the ability to make causative assumptions from response times alone. However, the strategic importance of response times for key stakeholders indicates a need to include these times as part of any measure of operational effectiveness.


The priorities are:

‘1. identify and assess the full range of foreseeable fire and rescue related risks their areas face,’
2. make appropriate provision for prevention and protection activities and response to fire and rescue related incidents

3. Collaborate with emergency services and other local and national partners to increase the efficiency and effectiveness of service provision

4. be accountable to communities for the service they provide, and

5. develop and maintain a workforce that is resilient, skilled, flexible and diverse.’

(Home Office, 2018a pp.4)

We argue that only with clear operational performance measures can the newly formed independent inspectorate provide the crucial assurance on the effectiveness and efficiency of fire and rescue services. Therefore, these should address Priority 4 through a demonstration of Priorities 1, 2, 3, and 5.

Murphy and Greenhalgh (2013) call for three principles that could be employed to map more rigorously and accurately FRS performance. Firstly, national and local indicators would allow for empirical comparisons over time and across services. These should be ‘quality assured, robust and accredited standards and benchmarks publicly available from an independent host’ (Murphy and Greenhalgh 2013, 231). Secondly, key documents such as the Integrated Risk Management Plans (IRMP) should be analysed to identify areas of importance to each FRS. Thirdly, the commissioning of academic and operational research will help move the FRS towards a more mature performance monitoring regime and stimulate dissemination of good practice; however, this can only be achieved once there are clear measures to allow for meaningful comparisons to be made. Within public services, multi-dimensional performance assessments provide a valuable method to frame comparative and
temporal analyses with a call to test the legitimacy, coherence and durability of performance regimes (Martin, Nutley, Downe and Grace 2016).

Conventional operations management (OM) performance objectives focus on quality, cost, flexibility, speed, and dependability (Slack, Chambers, and Johnston 2010). Performance measures can then be designed to ascertain how closely an organisation is to meeting those objectives. Bateman, Maher, and Randall (2016) applied OM principles to measure fire service performance based on a development of Overall Equipment Effectiveness (OEE) (Nakajima 1984), which combines the operation, maintenance, and management of equipment to measure effective use of resources. Endorsement from the UK government (Department of Trade and Industry, 2000) has led to wide acceptance of OEE within UK manufacturing. It measures the effectiveness of equipment via the three indicators of performance ($p$), availability ($a$), and quality ($q$) and is used to track and trace improvements or decrements over time. OEE has been applied in the road freight industry (Simons, Mason, and Gardner 2004) and using OEE in the FRS involves viewing a fire appliance with its crew as a piece of operational equipment. Within the service industry the areas of availability, performance and quality are highly relevant to service delivery and planning is challenging in the emergency services (Naesens and Gelders 2005) and Bateman et al’s (2016) work with a UK FRS sufficiently positioned FRS performance measures within the OEE framework.

Overall operational effectiveness involves complex interactions between different processes and the isolation of data can miss these interactions when using OEE (Muchiri and Pintelon 2008). Priority 1 and 2 for the FRS is to ‘identify and assess the full range of foreseeable fire and rescue related risks their areas face,’ and ‘make provision for prevention and protection activities and respond to incidents appropriately.’ This suggests an interaction
between different areas of the service and more than simply the fire appliance with its crew.

When considering this alongside the suggestions from FRS stakeholders, we propose an alternative model assessing \textit{Responsiveness}, \textit{Availability}, and \textit{Preparedness} (see Table 1) to evaluate overall operational effectiveness in the FRS.

\textit{Table 1. Overview of Overall Effectiveness of Fire Operations (OEFO) Framework}

<table>
<thead>
<tr>
<th>Conventional Overall Equipment Effectiveness (OEE)</th>
<th>Overall Effectiveness of Fire Operations (OEFO)</th>
<th>Related Performance Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>Responsiveness</td>
<td>Turnout times</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Attendance times</td>
</tr>
<tr>
<td>Availability</td>
<td>Availability</td>
<td>Sickness (neg)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pump availability (neg)</td>
</tr>
<tr>
<td>Quality</td>
<td>Preparedness</td>
<td>Equipment maintenance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Training competency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Critical equipment fails (neg)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Breathing Apparatus tests</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commendations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vehicle accidents (neg)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discipline cases (neg)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Personal injuries (neg)</td>
</tr>
</tbody>
</table>

\textit{The Research Context: Demand, Capacity and Alternative Crewing}

Demand from all incidents experienced by the UK FRS has reduced by 46\% since 2002 (Department of Communities and Local Government 2013)$^2$. Thus, the Counties has explored more efficient ways of managing capacity including ACAs which allow fire stations to be staffed by whole-time firefighters but at a lower level of capacity than conventional

---

$^2$ Official statistics (Home Office 2017$^a$) report a 37\% increase in fire deaths in 2017 compared to the previous year, however these figures include the 71 fatalities from the Grenfell Tower fire. Anecdotally, the UK national press and the Fire Brigades Union (The Guardian 2017; Fire Brigades Union 2017) have reported an increase in fire related fatalities, however confirmation as to whether this is a continuing trend is not possible within the time frame of this paper.

(2:2:4) fire stations. ACAs offer a potential alternative to reducing capacity without a reduction in responsiveness along with a cost saving of up to £370,000 per year per station (The Counties 2012). In 2002 there were 31,761 wholetime (FTE) firefighters in the UK but this reduced by approximately 27% (22,957) by 2017 (Home Office 2018). Over the same period the number of retained firefighters (FTE) grew between 2002 and 2011 but by 2017 numbers had fallen to below that of 2002 (10,092) (Home Office 2018).

Reduction in demand for emergency incidents offers a rationale for lowering capacity. However, public expectation is that the FRS will maintain and improve responsiveness to emergencies. Response times need to be maintained and geographic coverage must be upheld in line with Fire and Rescue Services Act (2004). Fewer emergency incidents alongside demands to maintain coverage and responsiveness present a challenge to FRSs when considering reducing staff. In manufacturing, capacity is usually linked to output. For services, capacity is more difficult to define and is a balance between human and physical resources. In service industries, fluctuating demand requires adjustments to capacity. Seasonality and variability of demand is a particular challenge for the emergency services where the ability to respond is paramount (Richie and Walley 2016).

Reform and efficiency measures within Fire and Rescue Services have been, in general, driven by financial concerns rather than in response to demand, with the recommendation for more creative ways to adapt to the changing nature of the fire service (Knight 2013). Projections of the impact of austerity measures on the FRS predicted a “perfect storm” in 2013-15 culminating in the loss of frontline services (Chief Fire Officers Association 2012, pp. 4) and many services at this time looked at innovative ways to reduce operational costs in line with local priorities without impacting on the service delivery of prevent, protect, respond (Knight, 2013).
The traditional model of crewing (two day shifts followed by two night shifts and then four days rest known colloquially as 2:2:4) was seen as one area open to reform. The changing proportions of fire safety and emergency response places more responsibility in the hands of firefighters, ensuring fire appliances remain crewed and available through self-rostering and annualised hours (Knight 2013). By matching the needs of the service against the resources available, a more efficient use of operational personnel can be deployed. Alternative Crewing Arrangements (ACA) is one such intervention, maximising the use of operational resources whilst minimising cost and ensuring public perception of service delivery remains relatively unchanged (The Counties 2012a)3.

The Counties ACA involved firefighters living in stations for up to five continuous days and nights, with 12-hours ‘down time’ in each 24-hours. The ACA required 50% fewer personnel than traditional 2:2:4 shift patterns but with no reduction in geographic cover (The Counties 2012a) and is an attractive crewing arrangement for fire stations that have lower call-out demands. A 2:2:4 system requires seven personnel to crew a ‘watch’ on a one pump (fire engine) wholetime station, including the Junior Officers (JOs), with four watches per station; a total of 28 operational crew. By contrast, ACA requires 14 crew using a self-rostering system to provide the same level of cover, providing a total reduction in operating costs of £370,000 per station per annum (The Counties 2012a). An overview of the key structural differences between ACA and the traditional shift system can be found in Table 2.

This paper forms one part of a much wider mixed-methods study investigating the impact of the change in shift system on both operational performance and employee wellbeing. Wellbeing data and overall findings from the whole project are reported elsewhere.

(Maher 2018, 2019). The scope of the current paper is to explore the utility of an objective framework to benchmark operational performance in the FRS and track changes over time in order to draw conclusions on operational outcomes of the change. The Overall Effectiveness of Fire Operations framework used in this study is outlined in Table 1 and was adapted from Bateman, Maher, and Randall (2016) to provide a more appropriate framework for the evaluation of ACA. Quality was translated to preparedness with the focus towards the organisation being prepared to meet the demands of its core business activity (attending and dealing with emergency incidents). It was broadened to include more than just personnel measures, e.g. equipment failures.

Table 2. Comparison of ACA with typical 2:2:4 shift system

<table>
<thead>
<tr>
<th>ACA</th>
<th>2:2:4</th>
</tr>
</thead>
<tbody>
<tr>
<td>24-hour shifts</td>
<td>12-hour shifts</td>
</tr>
<tr>
<td>Self-rostering of up to five consecutive 24-hour shifts in one batch</td>
<td>Continuously rolling shift pattern</td>
</tr>
<tr>
<td>Shift broken into 12 positive, active work hours and 12 negative rest hours</td>
<td>All time at work is allocated as positive active work hours</td>
</tr>
<tr>
<td>7pm start time for all shifts</td>
<td>7am start time for day shifts and 7pm start time for night shifts</td>
</tr>
<tr>
<td>One crew of 14-16 personnel</td>
<td>Four watches of seven or more personnel</td>
</tr>
<tr>
<td>Top heavy Junior Officer structure (two watch managers and two crew managers per crew)</td>
<td>One watch manager and one (or two for a larger stations) crew managers per crew/watch</td>
</tr>
<tr>
<td>27% enhancement on basic salary as remuneration for being available for emergency calls during negative hours. Enhancement also included in pension contributions</td>
<td>Basic salary</td>
</tr>
<tr>
<td>References/station duties devolved down to firefighter level</td>
<td>Responsibility for references and station duties falls to watch and crew managers</td>
</tr>
<tr>
<td>Private, en-suite accommodation provided for rest during negative hours</td>
<td>Communal dormitories provided for rest during night shifts</td>
</tr>
<tr>
<td>Official family-friendly policy allowing families to visit the fire station during negative hours</td>
<td>No official allowance for families to visit the fire station</td>
</tr>
</tbody>
</table>

# exact number of personnel is dependent upon the number of appliances and specialist equipment at each station
Method

Pilot study

We used a pilot study to identify and explore the available data collected by The Counties FRS in line with the framework above for two reasons; 1) to assess the data’s suitability for the evaluation of the impact of ACA prior to information from other stations being gathered and interrogated, and 2) to identify a suitable timeframe over which data was to be collect to ensure sufficient yield. The pilot study allowed us to make refinements to the OEFO framework for the full study. The Counties FRS collects data on various Key Performance Indicators using different computer-based systems. To streamline data handling, two members of specialist staff from within the service with responsibility for statutory data reporting provided us with a single document of merged data from the various sources. As part of the evaluation of The Counties operational effectiveness following the change to the new crewing system, a list of current performance measures was compiled through discussion with these members of staff and follows similar measures of a comparable FRS (Bateman, Maher, and Randall 2016).

The pilot comprises data for all identified measures within the OEFO framework for 2012 (pre-ACA) and 2013 (post-ACA) for Station D within a three-month period. This timeframe was initially considered by the data analysts within The Counties to provide sufficient data to allow for meaningful comparisons pre- and post- ACA. A month by month breakdown of performance data is not currently publicly available for the Counties which posed problematic for selecting an appropriate three-month period. Following advice from the data analysts at The Counties FRS the months May-July were selected because, historically, this period would be the busiest in terms of operational demand and would yield the most data.

Pilot Study Findings
Responsiveness. A three-month data collection period was not long enough to gather valid findings for the responsiveness domain as there were only 10 life critical incidents in the three-month period in 2012 and 9 in 2013. A longer time period (12-months) for both pre- and post- ACA is to be examined in the main study.

Availability. The availability of the wholetime appliance was very high at 99.5% or above. The strategic importance of wholetime appliances, may require extra resources obtained to keep these appliances available. Factors influencing the availability of appliances and any differences for retained versus wholetime could identify any impact from ACA in the full study.

Preparedness. Preparedness was assessed looking at a range of measures such as firefighter training, equipment testing and community events. Bateman, Maher, and Randall (2016) identified these measures as accessing the ability for firefighters to respond as expected to emergency incidents, alongside value added activities within the community. The pilot study data is outlined in Table 3.

<table>
<thead>
<tr>
<th>Assessment Measure</th>
<th>Pre-ACA</th>
<th>Post-ACA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equipment testing (N° of tests)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Weekly</td>
<td>433</td>
<td>506</td>
</tr>
<tr>
<td>• Monthly</td>
<td>16</td>
<td>65</td>
</tr>
<tr>
<td>• Quarterly</td>
<td>68</td>
<td>60</td>
</tr>
<tr>
<td>Hydrant checks (N° of checks)</td>
<td>110</td>
<td>78</td>
</tr>
<tr>
<td>Community events (time in minutes)</td>
<td>6840</td>
<td>9680</td>
</tr>
</tbody>
</table>
Feedback from data analysts at the Counties FRS on the utility of this data indicated that direct before and after comparisons over a three-month period would be crude. Hydrant checks needed to be assessed over a longer period; a 24-month testing cycle creates peaks and troughs over the course of the cycle as this work is conducted around other priorities. The number of visits to properties for Home Fire Safety Checks (HFSC’s) needs to be considered alongside successful completion of such visits due to occupants not being available and the need to revisit. Referrals from other agencies for HFSCs can also influence how many are attempted and completed. Procedural changes unrelated to ACA, such as the introduction of Redkite (safety management system for testing of operational equipment) and the use of equipment in operational situations may impact on equipment testing statistics. Other factors identified were either not available, infrequent events, or partially represented in another measure, for example personal injuries would show in sickness absence in availability.

**Conclusion from Pilot study**

The three domains of the OEFO framework (responsiveness, availability and preparedness) provide a useful measure of performance following the introduction of ACA. The final framework for assessing the operational effectiveness is summarised in Table 4, and includes the performance indicators identified from the pilot study as having the greatest utility.

**Table 4. OEFO Framework for full study**

<table>
<thead>
<tr>
<th>Overall Effectiveness of Fire Operations (OEFO) category</th>
<th>Key performance measures (unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responsiveness</strong></td>
<td>Turnout times (sec)</td>
</tr>
<tr>
<td></td>
<td>Attendance times (sec)</td>
</tr>
<tr>
<td><strong>Availability</strong></td>
<td>Sickness (days/person)</td>
</tr>
<tr>
<td></td>
<td>Pump availability (%)</td>
</tr>
<tr>
<td><strong>Preparedness</strong></td>
<td>Training drills (n°/person)</td>
</tr>
<tr>
<td></td>
<td>HFSC’s (n°/station)</td>
</tr>
</tbody>
</table>
Following assessment of the available data, three months did not provide a great enough yield of data to be able to infer impact of ACA on performance. Therefore, the full study draws on 12-months data pre and post ACA for each measure.

**Full Study**

**Data collection method**

12-months existing performance data for each of the measures in Table 4 were collected for each station pre- and post-ACA starting from each station’s respective date of ACA implementation. With Station A being a new station and therefore having no pre-ACA data, data for all the Counties FRS wholetime stations in combination were collected to provide an overall mean for the service as a comparison.

The same two members of the Counties FRS staff as for the pilot study were responsible for collating the data from the relevant service databases and collating it within an excel spreadsheet for the principle investigator (PI). Each measure’s data were presented as means per month, per station. Measurement of performance is at the station level, both before and after implementation of the new shift system allowing for the nuances of ACA introduction at individual stations to be explored with the context of each in mind. Station level performance data can then be compared to the service as a whole, providing an interpretative framework especially for new fire stations where pre-ACA data is unavailable.

**Analysis Method**

Descriptive statistics for each of the measures were scrutinised for percentage changes between the two time-points. For explanatory purposes a difference of 0-3% was identified as
minimal/no change, 4-10% was identified as a small change (positive or negative) and above 10% was identified as a large change (positive or negative). The purpose of an initial screen of percentage change was to identify any areas that may be problematic for service users in terms of practical significance but may not show statistical significance. For example, an increase in mean response time by 30 seconds may not be statistically significant but could have serious consequences at an emergency incident.

Findings have been categorized as per the OEFO framework (Table 4); Responsiveness, Availability and Preparedness. Examination of the assumptions for parametric testing revealed the standard error of skew was acceptable (<1.96) and so t-tests were used to assess whether any differences are statistically significant to identify any change that could be potentially attributed to ACA. Repeated measures t-tests were performed for Station B, Station C, and Station D, where before and after data are available. Independent samples t-tests were performed for Station A against The Counties means, as pre-ACA data is unavailable for this station. Due to the number of t-tests performed a Bonferroni correction was applied because there is the potential for findings to guide operational decisions within The Counties FRS, thus a more conservative approach is justified (Armstrong, 2014). Statistical significance was accepted at $p \leq .002$, and the magnitude of any effect was determined by using a Pearson’s correlation coefficient ($r$) using Cohen’s (2003) benchmarks (small effect $\geq .2$, medium effect $\geq .4$ and large effect $\geq .6$).

It is important to note that a change may not be statistically significant but still be of importance to the Counties FRS. Small changes that fail to reach statistical significance have not been dismissed if it falls below the acceptable standard for the Counties FRS. Findings are discussed in line with targets laid down in the Counties Integrated Risk Management Plan.
(IRMP) and other strategic documents to ensure relevance to stakeholders and within the boundaries of acceptable change.

**Findings**

**Responsiveness**

For life critical incidents (Table 6), Station C has remained stable for turnout times from pre- to post-ACA. Station A appears in line with the Counties mean for wholetime stations. The average turn-out time for Station B and Station D has increased (i.e. was slower) by 9% for both stations. Comparisons were non-significant for all stations.

Attendance times (time mobile to time at scene) for life critical incidents have a mixed picture. Station D has a reduced average attendance time, Station C has increased, and Station B has remained stable. Station A has a slower than average attendance when compared to the service average. The percentage of calls achieved within the Counties target time of under 10 minutes (The Counties 2012) remained high and showed improvement for Station B and Station C. Station A shows a higher percentage of life critical incidents attended within the same target time when compared to the Counties mean. However, Station D shows a moderate decline in the percentage attended within 10 minutes. The comparisons were again non-significant for all stations.)

Table 5. Results for responsiveness measures

### Life critical incidents (turnout time, attendance time)

<table>
<thead>
<tr>
<th>Counties mean</th>
<th>Mean N° Life critical calls per month</th>
<th>Mean turnout time (sec) (SD)</th>
<th>Mean attendance time (sec) (SD)</th>
<th>% Change</th>
<th>t (DoF)</th>
<th>p</th>
<th>% on target</th>
<th>% on target</th>
<th>% change</th>
<th>t (DoF)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre ACA Post ACA</td>
<td>Pre ACA Post ACA % Change t (DoF) p</td>
<td>Pre ACA % on target post ACA % on target % change t (DoF) p</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counties</td>
<td>3.81 n/a</td>
<td>151.75 (4.94) 145.92 (7.80) -2% -2.19 (18.61) .042</td>
<td>320.33 (17.33) 95.92 n/a</td>
<td>360.17 (76.42) 98.33 n/a 15% 1.76 (12.13) .103</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Station A</td>
<td>n/a 3.17</td>
<td>n/a 160.33 (18.68) 175.58 (24.94) 9% -1.66 (11) .125</td>
<td>431.83 (50.05) 92.75 n/a</td>
<td>453.75 (58.19) 93.08 5% -904 (11) .385</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Station B</td>
<td>3.5 3.41</td>
<td>179.42 (13.82) 179.75 (17.78) 2% -1.19 (11) .259</td>
<td>351.75 (32.64) 94</td>
<td>378.25 (25.78) 96.75 7% -2.42 (11) .034</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Station C</td>
<td>4.25 3.91</td>
<td>164.25 (23.11) 179.75 (13.54) 9% -1.75 (11) .108</td>
<td>505.33 (74.92) 82.83</td>
<td>461.92 (59.87) 75.25 -6% 1.62 (11) .133</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Station D</td>
<td>2.25 2.75</td>
<td>n/a 23.17 98.83</td>
<td>n/a 28.75 99.42</td>
<td>n/a 30.09 99.73</td>
<td>n/a 17.33 99.5</td>
<td>n/a -1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Non-life risk responsiveness (% under 20 min target time)

<table>
<thead>
<tr>
<th>Counties mean</th>
<th>Mean N° Life critical calls per month (pre ACA)</th>
<th>% attended within target time</th>
<th>Mean N° Life critical calls per month (post ACA)</th>
<th>% attended within target time</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29.9 n/a</td>
<td>99.42</td>
<td>n/a</td>
<td>99.42</td>
<td>n/a</td>
</tr>
<tr>
<td>Station A</td>
<td>n/a</td>
<td>n/a</td>
<td>23.17 98.83</td>
<td>99.42 99.73</td>
<td>-1%</td>
</tr>
<tr>
<td>Station B</td>
<td>28.92</td>
<td>99.75</td>
<td>28.75 99.42</td>
<td>99.73 99.5</td>
<td>-1%</td>
</tr>
<tr>
<td>Station C</td>
<td>35</td>
<td>98.92</td>
<td>30.09 99.73</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Station D</td>
<td>16.33</td>
<td>99</td>
<td>17.33 99.5</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
The responsiveness data for attendance time to non-life critical incidents (normal road speed calls) indicated that percentage attendance within the Counties target of 20 minutes (The Counties 2012) has remained very high, with a minimum of 98.92% pre-ACA and 98.83% post-ACA. ACA seems to have had little effect on attendance times for these types of incidents. The narrow variability in the data suggests this measure lacks the sensitivity to uncover any influence and no further analysis was conducted.

**Availability**

A summary of the availability data can be found in Table 7. The availability of the wholetime appliance was very high in both periods (98.56% or above) at all stations. There is a small reduction at all stations between pre- and post-ACA however this is small and potentially due to natural fluctuation. A more significant reduction is seen in retained appliance availability at the stations under investigation and across the wider service. This is most marked at Station D. Comparisons for wholetime availability for all stations were non-significant.

Retained availability was not significantly different between the two-time frames at Station B. Retained availability was significantly reduced in the 12-months after ACA for Station C (Mean pre-ACA = 98.51, Mean post-ACA = 95.08, $t(11) = 4.04, p = .002, r = .773$), Station D (Mean pre-ACA = 93.49, Mean post-ACA = 63.25, $t(11) = 10.64, p < .001, r = .955$) and the Counties overall retained complement (Mean pre-ACA = 92.74, Mean post-ACA = 89.21, $t(11) = 8.97, p < .001, r = .938$). Large effect sizes were seen for all comparisons ($r \geq .5$). No retained appliance operates at Station A.

There is a large percentage reduction in sickness absence for Station C and D. Station B saw a large percentage increase in the same measure. There is a large percentage difference
in sickness absence between Station A and the Counties mean. All comparisons were not statistically significant at the corrected value.

**Table 6. Results from availability and preparedness measures**

<table>
<thead>
<tr>
<th>Availability (wholetime and retained appliance, and sickness absence)</th>
<th>WT pump availability</th>
<th>Retained pump availability</th>
<th>Sickness absence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre ACA</td>
<td>Post ACA</td>
<td>% Change</td>
</tr>
<tr>
<td>Counties mean</td>
<td>99.45</td>
<td>99.24</td>
<td>-0.21</td>
</tr>
<tr>
<td>Station A</td>
<td>n/a</td>
<td>99.28</td>
<td>-0.17</td>
</tr>
<tr>
<td>Station B</td>
<td>99.75</td>
<td>98.56</td>
<td>-1.19</td>
</tr>
<tr>
<td>Station C</td>
<td>99.41</td>
<td>98.97</td>
<td>-0.44</td>
</tr>
<tr>
<td>Station D</td>
<td>99.71</td>
<td>99.63</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

| Preparedness (training drills and HFSC's) | | | |
| Mean N° drills per station per month | Mean N° drills per person per month | Mean N°HFSC's per month | | | | | | | | | | | |
| | Pre ACA | Post ACA | % Change | Pre ACA | Post ACA | % Change | t (DoF) | p | Pre ACA | Post ACA | % Change | t (DoF) | p |
| Counties mean | 35.96 | n/a | n/a | n/a | n/a | n/a | 36.68 | n/a | n/a | n/a | n/a | n/a | |
| Station A | n/a | 13.17 | -63 | n/a | 11.29 | n/a | 27.67 | -25 | n/a | n/a | n/a | | |
| Station B | 31.67 | 16.92 | -47 | 13.57 | 14.5 | 7 | -7.31 (11) | 480 | 25 | 25.83 | 3 | 1.125 (11) | .903 |
| Station C | 35.67 | 14.08 | -61 | 15.29 | 12.07 | -21 | 1.38 (11) | .195 | 34.17 | 33.25 | -3 | 1.133 (11) | .897 |
| Station D | 28.83 | 20.05 | -29 | 12.36 | 17.64 | 43 | -3.02 (11) | .012 | 7.33 | 4.58 | -38 | 1.42 (11) | .183 |
Preparedness

The yield of data for measures of critical incidents, commendations and personal injuries were not high enough for meaningful comparisons, omitting them from the analysis. Data for equipment testing has very little variability due to being a routine action incorporated into ACA activities, and so was also omitted from the analysis. Equivalent before and after comparisons for hydrant checks are not possible due to increasing numbers being performed by non-operational personnel.

The greatest variability was found within data for training drills and Home Fire Safety Checks (HFSC). A summary of the data is found in Table 7.

The number of drills performed by the station appears to have had a large decline but this does not account for fewer personnel on the station following the introduction of ACA. To control the number of firefighters the mean number of drills per person per year were calculated, to allow for a meaningful comparison between time periods. Once this measure was applied, the number of drills carried out before and after ACA remained relatively stable. Station D is an exception, where personnel carried out more drills per year in the 12-months following ACA introduction. None of the comparisons within this measure reached statistical significance.

HFSC’s have remained stable for Station B and Station C since the introduction of ACA. Station D appears to show a decline in the number performed, however, the low starting point means that this difference is minimal. The number of HFSC’s performed by Station A were unable to be compared to The Counties means as data included HFSC’s carried out by non-operational personnel thus data are not comparable. The remaining comparisons were not found to be statistically significant.
Discussion

Our aim for this paper was to present a framework to measure operational effectiveness in the Fire and Rescue Service in order to provide a commentary on the impact of an alternative crewing system on service delivery. Utilising established approaches, the OEFO organised existing performance data in a meaningful way to allow comparisons pre- and post-ACA. The relative infrequency of some FRS activities, e.g. health and safety events, and insensitivity of some measures required data to be drawn from a variety of sources.

Alternative crewing and operational effectiveness

Overall, ACA was found to have mixed effects on operational effectiveness but no major positive or negative effects at any of the stations under investigation. Some areas may be more sensitive to this new working arrangement than others, such as retained appliance availability. However, there are caveats and The Counties FRS and other services nationally need to consider wider implications before rolling out the system more extensively.

The Counties FRS introduced ACA into stations with a low call profile, reducing the chances of regular disturbance during negative hours and increasing the likelihood of longer periods of active work during the day time. Positive hours allow time for the maintenance of core skills and equipment, and to carry out statutory fire safety duties on which stations are targeted. Therefore, stations with higher call profiles, particularly during the night time, may have difficulty in completing all the tasks required. The OEFO can provide stations with the opportunity to identify bottlenecks or deficiencies in performance indicators in a timely manner to prevent a negative impact on service delivery, similar to that used within the manufacturing sector (Bateman, Philip, and Warrender 2017).

Responsiveness

Turnout times for life-critical calls for Station A and Station C have remained stable, whereas Station B and Station D have seen moderately slower times in the 12-months
following ACA. These moderate differences were not statistically significant; however, for Station D the mean turn out time in the 12-months post ACA did fall above the Counties critical target of three minutes (The Counties 2012a).

Findings for life-critical attendance times are varied. Two of the existing stations, B and C, have seen a moderate increase, but the other existing station, D, has seen a moderate decrease in attendance times. The large difference between Station A and the Counties mean can be attributed to the differences in station ground; the area of coverage offered by each station. The Counties mean includes city fire stations which have a much smaller coverage and attendance times at city stations are well within the 10-minutes and quicker than rural stations, such as Station A. The yearly means for all stations falls below the Counties target of under 10 minutes (The Counties 2012b). However, at Station D the percentage of calls attended within the target time dropped by 7%. Station D covers a large rural area, and so other variables, including distance of the incident from the station and weather conditions reducing the speed of the appliance, have greater influence. Whilst these factors apply for the other three stations, Station D covers 154 square miles compared to the next largest at 107 square miles. Attendance times have the greatest potential of interference from extraneous variables and the idiosyncratic nature of the stations make assumptions about the influence of ACA difficult. As part of the data gathered for each incident, any records given for attendance slower than 10 minutes would be particularly useful and could be examined for factors that may be linked to ACA.

Attendance at non-life-risk (normal road speed) incidents within the target (20 minutes) remains very high, above 98%, before and after ACA across all stations. The narrow variability in this data limits this measure’s usefulness in assessing responsiveness. The data provided by the Counties was in the form of the number of non-life-risk calls and the
percentage attended within target per station per month. Scrutiny of actual times attended in seconds would allow for comparable analysis to the life-critical incidents and increase the sensitivity of the data; the target time of 20 minutes may mask variability across stations and across time.

Response times (turnout and attendance) appear to provide a useful measure of responsiveness, but the findings support the relative insensitivity of this data from a lack of variability (Bateman, Maher, and Randall 2016; Carvalho et al. 2007; Jaldell 2005). The low number of life-critical calls in the data set reduces variability and the statistical power that can be provided by measures of responsiveness. As such the findings support the need to assess additional performance measures to provide a fuller picture of operational effectiveness.

Availability
Availability has had the most marked change following the introduction of ACA. Wholetime appliance availability has remained high across all stations. The strategic importance of these appliances may motivate the service to move resources, including firefighters, to keep appliances available.

Retained appliance availability has seen a large percentage reduction across all the stations and the service as a whole. For two of the stations, C and D, and for the Counties’ entire retained complement this reduction was statistically significant, and Station B was approaching the corrected significance value. Interviews revealed that there was a reduction in hours ACA staff were able to offer at retained stations as part of dual-contract, and dual-contract personnel ended retained commitments once starting ACA. The wider impact on retained availability should be considered by the FRS when implementing ACA, including non-ACA stations with a reliance on dual-contract personnel.
Sickness has seen a large improvement across two stations, C and D, as well as a large percentage difference between Station A and the Counties mean. These differences failed to reach significance once the correction was applied, however, the low starting point may mask an effect. From the last available reports capturing sickness absence data across the UK, FRS personnel were shown to have a higher than average levels of sickness absence when compared to other workers (6.3 vs 4.9 days per person per year) (Black and Frost 2011; Department for Communities and Local Government 2010). The Counties sickness absence rate of 4.4 days per person per year compare favourably to the UK FRS as a whole, potentially due to a healthy worker effect following the introduction of mandatory health screening, fitness tests, and flu inoculations for operational personnel (as stated by Occupational Health Manager, and Health and Fitness Advisor for the Counties 2013); thus, the large percentage change is still worthy of further investigation. Discussions with station personnel revealed informal arrangements cover through swapping of shifts rather than a lower incidence of ill health. Station B saw a large percentage increase in sickness absence. Following interrogation of the sickness absence data supplied, this increase in the mean appears to be due to one period of long-term sickness by one crew member skewing the data. The low starting point of this data means one extended period of sickness may influence the overall results and therefore no conclusions can be drawn as to whether Station B follows a different trend to the others.

**Preparedness**

Following the pilot study, procedural changes in the wider service impacted on post ACA data. Increasing numbers of hydrant checks are performed by specific non-operational personnel, reducing the need for stations to undertake these checks. Any difference between the two time-points cannot be attributed to ACA because of this procedural change.
The number of drills performed per person per month may be influenced by the introduction of ACA because the 12 hours down-time reduces the time available for training. Station C saw a large percentage decrease in the number of drills performed. The Bonferroni correction applied to the analysis may have led to a type II error for this variable, due to the conservative nature of the correction, thus reducing the power to uncover an effect (Gelman, Hill, and Yajima 2012). Interrogation of the raw emergency call data could uncover whether the time of day that Station C was called out has impacted on time available for training, through pushing back the start of the working day. By contrast, Station D has seen a large percentage increase in the number of drills performed following the introduction of ACA, which fell below the .05 significance level but failed to reach significance when the Bonferroni correction was applied. The low call profile for this station may be a factor in creating a greater proportion of time for training. The self-rostering nature of the shifts required risk critical training to be repeated more often to ensure all personnel maintain competency, potentially increasing the number of drills per person.

Training data was only available via station records, not for centralised training from the Training and Development Department. To maintain competency in key areas, breathing apparatus and live fire training are carried out centrally, involving the fire appliance from the station attending a training site with all the crew available that day. When using a watch-based system, this would entail visiting the training centre together as one watch with all crew trained simultaneously. With the self-rostering on ACA the same personnel rarely work together on a continuous basis, thus requiring multiple visits to the training centre to cover the same training. Training department data could be useful in quantifying whether this had any impact on the efficiency of the ACA stations through repeated visits to centralised training.
HFSCs are carried out in response to demographic information via the MOSAIC profiling system (Local Government Association 2012) to identify those most vulnerable to the risk of fire. Stations are responsible for visiting those households to offer advice and fit smoke alarms where necessary. Stations B and C saw a minimal change in the number of checks performed pre- and post-ACA. One explanation for this could be that HFSCs are a targeted indicator, so this work may be prioritised over non-target driven work, such as training. Station D had a large percentage decrease in the number of HFSCs performed, however, a low starting point means that in real terms this difference is minor. The demographics of the Station D area may explain the lower figure of HFSCs when compared to the other stations in the analysis. There are much lower levels of social deprivation and higher than average employment levels particularly in professional and managerial occupations (District Council Statistics 2015). As there are fewer vulnerable members of the community within this station area, this data will be more sensitive to change.

**Framing operational effectiveness in the Fire Service**

With operational effectiveness not clearly defined within the literature and the current performance management system in place in the FRS open to wide interpretation (Downe, Martin, and Doering 2017), a transparent and objective way to benchmark fire and rescue services is needed (Murphy and Greenhalgh 2013). Murphy and Greenhalgh (2013) also argue that academic and operational research on performance management within the FRS will help move it towards a mature monitoring regime in line with other public services. The conceptualisation of operational effectiveness for the FRS, as outlined above, uses key metrics of importance to the service providing both academics and practitioners with clear measures to make meaningful comparisons.

---

4 Reference anonymised for maintenance of The Counties FRS anonymity
The move towards an independent inspectorate has led to the consultation on a new Fire and Rescue Framework for England, placing an emphasis on transparency and accountability in line with Integrated Risk Management Plans and efficiency plans (Home Office 2017). The OEFO framework developed in this paper offers an innovative tool for assessing operational effectiveness in the FRS and could inform the inspectorate as to appropriate indicators for benchmarking fire and rescue services.

The Home Secretary expressed a desire for greater transparency and accountability for performance in the FRS (May 2016). Following the recommendations by Murphy and Greenhalgh (2013) on the use of national and local indicators, and commissioning of academic and operational research, the OEFO provides a framework to benchmark operational effectiveness as part of any new inspectorate and provide a base upon which empirical work can be formed. Effective performance centres around what is measured, and as previously discussed, current systems of benchmarking performance within the FRS do not provide clear comparative elements. The OEFO is the first tool to attempt to operationalise operational effectiveness within the FRS using an established and well researched aggregate measure from lean methodology. Adapting OEE for use outside of manufacturing has previously produced a robust and relevant measure for the sector to which it is applied (Simons, Mason, and Gardner 2004), supporting the feasibility of the OEFO framework.

Through analysis of the data routinely collected by The Counties FRS, some measures were found to map well on to the framework and offered clear comparisons before and after the change to ACA. Responsiveness and availability data have continuity of collection methods within the service and so there is confidence in the comparisons made. Preparedness measures are more open to interpretation on collection and are influenced by procedural changes within the service. An example of this would be the completion of hydrant checks,
increasingly carried out by non-operational personnel therefore a reduction in the number of checks may not be indicative of any influence by ACA. However, if this procedural change occurred in order to reduce the demand on ACA stations then it could have had an impact.

The OEFO framework has offered a useful foundation from which to explore the available performance data for The Counties FRS to test for any potential impact following the introduction of ACA at the four fire stations under investigation. The new framework was attained by translating the principles of OEE to measure the effective utilization of a fire appliance and its corresponding crew to the core business activity of attending and responding to emergency incidents. The categories across the two measures are, for all intents and purposes, the same, strengthening the validity of the OEFO.

The OEE and OVE tools benchmark actual performance against planned performance using a metric to provide overall percentage effectiveness. Due the difficulties in mapping planned performance for the core business activity in the FRS, through the variability in demand for emergency calls, creating an overall effectiveness percentage becomes problematic. Despite this, there is an opportunity to use operational effectiveness data via the OEFO framework regularly to monitor and manage station activity as part of continual improvement.

**Conclusion**

The development of the OEFO tool provides a rigorous approach for the tracking of fire service operations over time. The approach reflects the societal requirements for good response times but takes in the need for crews to operate effectively (and safely) whilst ensuring that preventative activities central to public safety are maintained. Our research considered how these aspects of fire service work can be exemplified through the three dimensions of responsiveness, availability, and preparedness. Whilst it is desirable to express
these in a single measure, this is likely to be over crude and mask potential problems. The assessment of public services beyond cost and simplistic single measures is crucial to service delivery (Bitici, Cavalieri, and Cieminski 2009). The OEFO allows for informed decision making and the monitoring of changes over time and between stations. It also ensures decision makers are held accountable by highlighting appropriate measures.

Using the OEFO tool, ACA was not found to have any overall positive or negative effect at the four stations under investigation. One of the aims of our project was to confirm whether ACA as an intervention, met its objectives of reducing operational costs without adversely affecting operational performance. ACA allows for the lowering of the total number of full-time firefighters and so the aim of no loss of coverage is largely substantiated. Wholetime operational effectiveness has not seen a negative impact overall and there have been improvements in sickness absence and training. However, a reduction in the depth of coverage may impact on the services ability to respond to major events that require a rotation of fire crews over a sustained period. Nationally accountable resources such as those for terrorist response, urban search and rescue, and hazardous materials require continuous availability under the terms of reference for the fire service (Civil Contingencies Act 2004) therefore are not suitable for ACA where there may be periods of mandatory rest for crew.

Services need to ensure such resources are housed within stations with suitably available crew at all times. Strategic planning for major incidents involves risk planning for a most likely scenario and extraordinary events such as the Grenfell Towers Fire or a terrorist incident such as 7/7 involve collaboration between services to draw on all available resources. The impact on retained availability needs further investigation to uncover the underlying mechanism of the negative result across the stations and the Counties as a whole, as this may also influence the services ability to respond to large scale or protracted incidents.
A direction for future work would be to explore ways to model operational effectiveness within strategic major incident planning to provide a quantifiable assessment of the wider impact of operational changes, such as ACA.

Effective performance centres on what is measured, and current systems of benchmarking performance within the FRS do not provide clear comparative elements. The OEFO provides a base upon which empirical work can be formed, as well as providing a practical framework for comparison both within and between services. The OEFO is the first tool to attempt to operationalise operational effectiveness within the FRS using an established and well-researched aggregate measure from operations management. Adapting the OEE for use outside of manufacturing has previously produced a robust and relevant measure for the sector to which it is applied (Simons, Mason, Gardner 2004), supporting the feasibility of the OEFO framework.

References


