East African Scholars Journal of Engineering and Computer Sciences
*Abbreviated Key Title: East African Scholars J Eng Comput Sci*
ISSN: 2617-4480 (Print) & ISSN: 2663-0346 (Online)
Published By East African Scholars Publisher, Kenya

OPEN ACCESS

Volume-5 | Issue-5 | Sept-2022 | DOI: 10.36349/easjecs.2022.v05i05.001

**Original Research Article**

# A Comprehensive Scoring System for Productivity Increases Engagement and Wellness Outcomes

Nitish Shrivastava[1*], Neel Tiwari[2], Elaiyaraja Thangavel[3]
[1]Senior Vice President, [2]Senior Product Manager, [3]Associate Vice President, Persistent Systems

**Article History**
Received: 15.08.2022
Accepted: 24.09.2022
Published: 29.09.2022

**Journal homepage:**
https://www.easpublisher.com

**Quick Response Code**

**Abstract:** The research conducted in this article was used to prove that there is an increase in employee engagement of productivity increasing activities when a scoring system is utilized by employees. Previous research conducted proved that it is possible to evolve a scoring system that takes into account not just activities tied to productivity, but also factors relevant (and tailored) to individuals based on bias, relevance, location, etc. In this research article, we show that when a set of subjects were asked to utilize a scoring system to grade activities conducted (with a control group of subjects who did not utilize a scoring system), the group using the scoring system were more likely to engage in activities that enhanced productivity compared to the control group.

**Keywords:** Work-life, productivity, health, wellness, flow, organizational resources, personal resources, positive, bias, relevance, score, formulae.



## INTRODUCTION

With the renewed focus on productivity and wellness, organizations the world over have recognized threats to productivity and wellness as fundamental threats to public health and safety as well, and that it has long term consequences for individuals and organizations alike. Numerous organizations consider employees as the main ingredient for plans to achieve financial goals, and believe that employee motivation is paramount to actualizing those goals (Saad, 2018). Most efforts revolve around evolving criteria that assesses performance of employees through some form of appraisal system; the driving motivation is to ensure that such appraisals are fair, provide an accurate assessment of employee strengths and weaknesses, and eliminate opportunities for insufficient awareness of goals (Bagul, 2013).

The ultimate goal, of course, is to increase employee engagement with respect to stated organization goals, and ensure that employees feel that they are being given sufficient resources, confidence and self-efficacy to maximize individual success on the path to these goals. As of now, most methods that organizations use to oversee and analyze employee engagement follow long-standing practices that assess factors like employee retention, level of daily attendance, perceived confidence and trust in the organization they belong to, and so on (Antony, 2018).

To foster this culture of high employee engagement, human resources departments in organizations employ a set of tools based on these drivers (Tomlinson, 2010). This becomes especially critical when dealing with teams that are scattered virtually across regions of the world, as is the norm. Some strategies involve utilizing theories like Job Demands-Resources Theory to assess drivers for engagement (Shaik, 2019).

Whatever be the case, it cannot be denied that organizational support is a critical piece of the puzzle in fostering employee engagement (Khajuria, Khan, 2022). And not only corporate organizations; organizations that rely on high volunteerism and perform public service also recognize the need to pour organization resources into initiatives that foster employee innovation (Knox, Marin-Cadavid, 2022).

The other side of the coin (detractions from employee engagement) is also not to be ignored. One of the factors that authors looked at was the effect of stress on employees, and its impact on productivity and

***Corresponding Author:** Nitish Shrivastava
Senior Vice President, Persistent Systems

wellness. Employee performance can be significantly degraded by physical and mental ailments that are caused by work life imbalance, lack of communication and general health (Bedarkar, Pandita, 2014). Research in the field of psychology suggests that its possible to increase employee engagement by stepping out of the traditional "one size fits all" approach and to instead focus on drivers of engagement tailored towards individuals (Crabb, 2011); it is even possible that such drivers can be actualized by use of coaching techniques tailored towards individuals and their experiences (Crabb, 2011); furthermore, efforts in this regard have been shown to not remain exclusive to employees, but can cause significant beneficial after-effects on the customers, which could be of particular importance to organizations in service eco-systems (Han, Chen, 2021). To this end, the authors' previous research attempted to consolidate these gathered observations and provide the genesis of a comprehensive approach towards fostering individual productivity and wellness.

Previous research by authors led them to two conclusions:

Conclusion a: Workplace stress was a significant factor in lowering productivity/wellness in working professionals; furthermore, a significant indicator of stress was an inability to keep up with deadlines, failure to deal with adverse events, loss of self-efficacy and lack of clarity on progress on workplace activities

Conclusion b: It is possible to represent the current state of productivity and wellness using a scoring system that takes into account factors like activities, time, profile, etc. that allows individuals to assess productivity/wellness along a scale that directly co-relates to productive activities undertaken by individuals.

The question that the authors were now seeking to answer was: what would be the practical applications of such a scoring system, and would it be possible to *prove* that individuals who utilized such a scoring system experienced better productivity and wellness outcomes compared to individuals who did not utilize this scoring system?

The following sections will first briefly recap the methodology used for the scoring system. It will then go into details about how the scoring system was utilized to *indicate* to utilizing individuals through *interventions* about the need to maximize specific activities; finally, it will outline how activities post-intervention were captured and analysed in order to prove a *significant change in productive activities occurred post-intervention*, thereby confirming the authors' hypothesis.

# MATERIALS AND METHODS

The authors of this research study relied on data gathered from a previous research study; the preceding study was built out of extensive research data that was collected from working professionals. We will first re-introduce materials and methods utilized for data from the preceding study, and then supplement additional information that explains our methodology for answering the relevant questions for *this* study.

## STATEMENT OF HYPOTHESIS

Null hypothesis (Ho)Time intervention analysis of engagement data from subjects that were prompted by our scoring system to improve scores showed no change in the mean level of the distribution of values indicating "current" engagement post prompt.

**Alternate hypothesis (Ha)**

Time intervention analysis of engagement data from subjects that were prompted by our scoring system to improve scores showed either: a) permanent constant change b) brief constant change c) gradual increase to the mean level of the distribution of values indicating "current" engagement post prompt.

## CONTROL VARIABLES EMPLOYED

In order to ensure widest spread of data, data was collected in the following proportions from subjects:

- Across multiple job disciplines (criteria explained in section *Organizing subjects based on profile* below)
- Equal numbers of men and women
- Equal proportions of shift-based employees (morning shift and night shift)
- Split in equal proportions across managerial/individual contributor roles
- Spanning multiple countries (United States, India)

## EVOLUTION OF METHODOLOGY

The next set of sections describes how the authors evolved the methodology to be employed for testing the hypothesis.

**Organizing subjects based on profile**

(Note: At the outset, we obtained explicit permission from our subjects to collect each data attribute that we utilized in our research).

Some of the information that we collected at the beginning was still useful. Since it was important that we have a wide spread of working professionals represented in our subject list, we identified some key "profiles" of working professionals that we wished to study and aligned them to specific jobs/professions to aid in classification and segregation. These included:

- Sales professionals
- Software developers
- Support engineers
- Product Managers
- QA engineers

It should be noted that the majority of these subjects were engaged over a period of *5 years* in terms of collection of data for our study.

**Evolving quantifiable categories of information to collect**

We spent some time interviewing subjects from each of these profiles. Instead of trying to rely on manual recorded observations, we first started with a set of specific questions:

- *What measures do you usually employ in your job profile to determine that your goals have been achieved?*
- *What conditions/outcomes during the course of your specific job profile would be considered as a failure to achieve your goals?*
- *Based on previous questions, how would you categorize your job performance?*

Asking these questions and arriving at a consensus resulted in a set of "criteria" for productivity/wellness for each job profile being evaluated.

**General wellness criteria among profiles**

For individuals belonging to these profiles, we intended to make a general case for analyzing wellness; accordingly, we captured certain common data across all profiles (like Heart rate, Sleep, Steps and Fitness activity)

**Sales professionals**

This profile corresponds to working professionals who work in sales to win deals that result in additional revenue for their employer.

For such individuals, success would be categorized under the following categories:

- Bringing in new sales leads that result in opportunities for increased revenue for the company
- Successfully closing deals with customers to realize additional revenue (and doing it as quickly as possible)

Conversely, there are a few scenarios that could be judged as a "failure" in productivity:

- Failure to bring in new leads over the course of a financial year
- Failure to close out deals, resulting in dropping these opportunities and preventing revenue from being realized
- Failure to *adequately* pursue open opportunities through available methods like customer in person meets, calls, emails/meetings, etc.

Armed with objectives for "gauging" productivity, we captured data from specific data sources:

- Sales related data (CRM)
- Emails/Meetings
- Travel information

**Software Developers**

This profile corresponds to software engineers who are directly or indirectly responsible for maintaining the code base of products/services/projects in an organization, whether it be by contributing to new features or fixing existing issues.

For such individuals, success would be categorized under the following categories:

- Timely contributions to the source code management system (which would imply quick closure of assigned defects, fast closure of requests for enhancements or new features)
- Good quality contributions (which would imply minimizing of defects arising from changes/fixes to the product, infrequent changes happening on touched source code files, etc.)

Failure in such cases in terms of productivity would include:

- Leaving open assigned defects opened for a long period of time without closure
- Bad quality of code contributions resulting in increased issues, and requiring more code rewrites, slowing down development, etc.

An attempt was made to capture data in these categories:

- Code check-ins in source code management systems utilized by subjects
- Bugs/feature requests in project management software
- Emails/Meetings
- Software App Usage

**Support Engineers**

This profile corresponds to working professionals who are responsible for directly interfacing with customers utilizing products/services from their organization (with the purpose of customer assistance/support, preliminary analysis, communication with backend teams, and closure of reported issues).

For such professionals, success criteria would include:

- Number of customer issues (aka "tickets") resolved
- Reduced time to resolve filed tickets

On the flip side, failure would include scenarios like:

- Taking too long to resolve tickets
- Having a higher number of critical/high priority tickets open without resolution

To account for such scenarios, we captured data in the following categories:

- Incident management system tickets
- Emails/Meetings

**Laying out patterns for a working professional's day**

To properly assess the data required to answer this question, it must be framed in a manner that can be tied to typical patterns of work and leisure that working professionals undergo. We started by taking a time interval of 1 day (24 hours) out of the life of a person; we can roughly categorize periods of the day in the manner below. Note that in this instance we are assuming that the subject works in the morning shift; in the case of working professionals who work in different shifts, the time periods and associated tasks/behaviours would change accordingly.

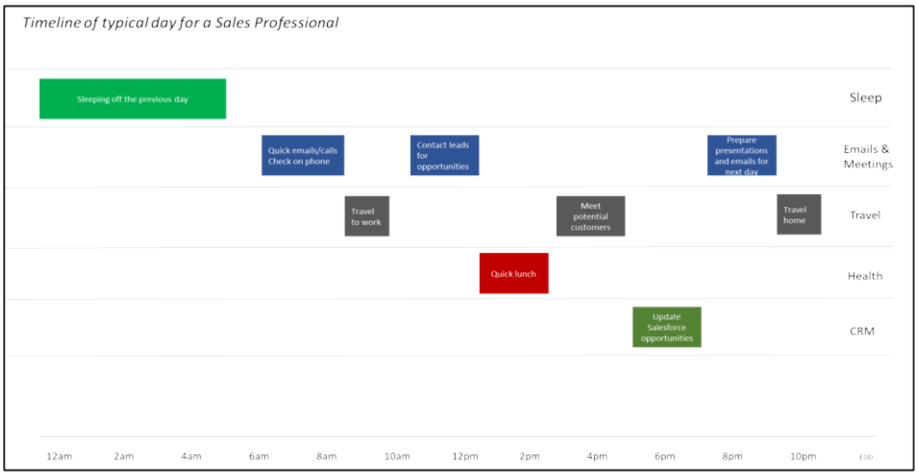
- *Sleep period:* This is the number of hours of the day during which a subject would typically be in a sleep state. The actual quality of sleep during this period would have to be judged by multiple factors (i.e., REM periods, number of times that subject woke up, amount of sleep, how closely it fit circadian rhythms, etc.).
- *Post sleep morning period:* Typically, this is the time just after the subject has woken up, where the subject would indulge in activities that would eventually transition into typical activities during the day; this could include time to brush, take a shower, morning constitutional, breakfast, etc. Note that it's possible that activities during this period may also include preparatory work for the rest of the day or may include physical activities from the point of view of exercise.
- *Office commute period:* This can vary from person to person (and may not even exist for a subject who works remotely 100% of the time). The time period can vary depending on the commute to work distance, condition of traffic based on time of day, etc.
- *Working hours:* This would be the period where a subject is *expected* to engage in most productive activities from the context of the job/profession.
- *Return commute period:* At the conclusion of the day, if the subject is working from an office, this period would coincide with the return journey back to subject's home.
- *Pre sleep period:* Usually associated with a "winding down" of the day (and can also include physical activities), including dinner and relaxation activities followed eventually by commencement of the sleep period.

At this point of the methodology, we had essentially constructed a "picture" of a person's day based on time periods of *presumed activity*. This point is crucial; even aside from the fact that these patterns of time periods can vary depending on "shifts" in which a working professional can operate, it also doesn't fully consider the *quality* of activities that are undertaken during these periods. We instinctively (which is to say, without the need for explicit measurement) can ascertain that it is very rare for a working professional to have periods of activity that occur with such consistency, and even in the event of said periods actually coinciding with the "expected" activities, it is rare that there not be some sort of interruption or negative effect on quality of the activity, be it physical, mental, psychological or otherwise.

The decision was then made to more finetune our picture of working professionals by relying on *sources of data* that are available throughout our subject's day.

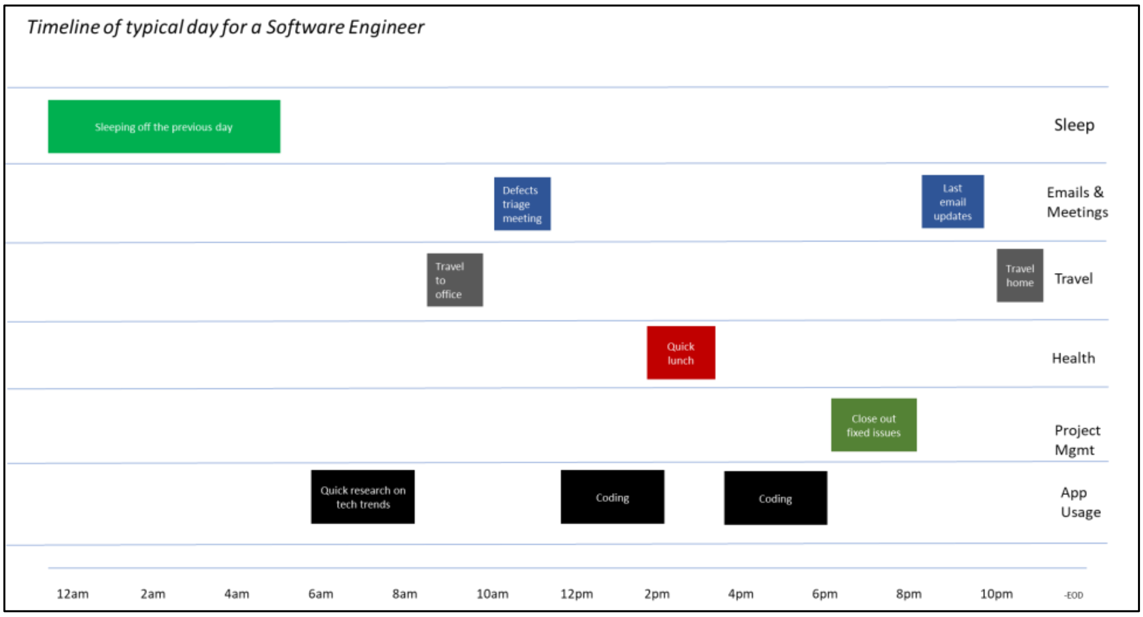
**Timeline for a Sales Professional**

To accomplish this, we interviewed sales professionals to build a "picture" of the sales professional's day. To do this, we create a timeline that models all 24 hours of a person's day, and then placing (based on their feedback) typical periods of activity. Accordingly, we come up with the following diagram for a particular sales professional:



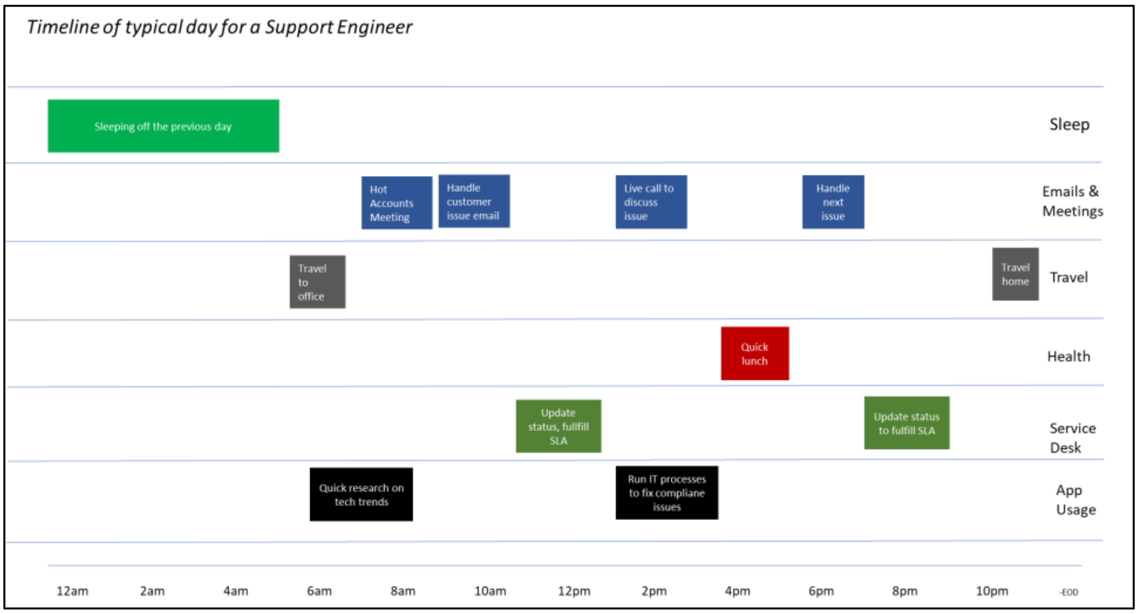
**Timeline for a Software Developer**

In a similar fashion, we came up with a representative timeline for software developers:




**Timeline for a Support Engineer**

A similar timeline was created for support engineers as follows:




**Organizing the set of subjects and data capture**

As we mentioned before, our original plan was to have our subjects maintain work diaries that they would write into over time. There were several problems with this approach; for one thing, it wasn't very reliable as a *comprehensive* record of activities since it depended on frequency and accuracy of written entries; for another, it could potentially *detract* from the efficiency with which work activities were undertaken, thus potentially undermining the study.

Once we had this realization, we realized that our only recourse would be to *automate* the collection of data from the subjects. Accordingly, we went back to the subjects, and determined what productivity tools were utilized for our subjects to do their day-to-day jobs? We gathered the answers to this question across all profiles and came up with a list of services/information to be gathered. Over time, we researched methods for gathering this information in an automated fashion (web services, APIs, software/apps, etc.), and began the process of monitoring and collection of information from subjects.

Accordingly, with appropriate disclosures of our intentions and with explicit permission obtained, we arranged for the capture of data from multiple subjects, using appropriate data sources to feed into our research. To briefly summarize the extent of data capture from subjects:

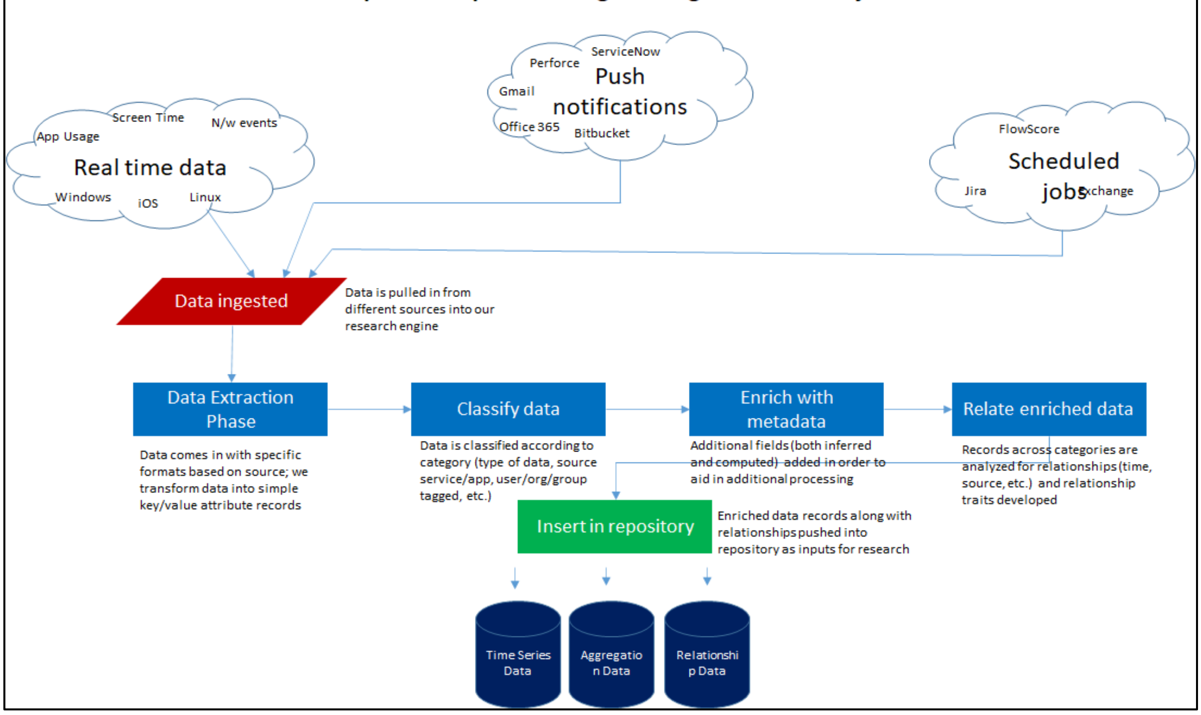
- **Health information:** We collected information about sleep, heartrate, fitness activities, etc. In addition, we also sync a health *score* that fitness tracker tools that subjects make use of.
- **Emails/meetings:** We collected emails/meetings information from popular office suites like Microsoft Office 365, Microsoft Exchange, Gmail, etc. with subject's consent.
- **Business desktop apps:** We collected information about screen time (defined as time spent actively working on a screen of a desktop PC, laptop, mobile, etc.), actual software/processes that the subject was working on, and factors that can influence degradation of work undertaken using business apps (i.e., network interruptions, machine restarts, etc.).
- **Business critical services:** We collected information from services that are utilized in some form or the other by subjects for specific purposes, i.e., CRM data from Salesforce, service desk tickets, tasks/issues filed in project management systems.
- **Mobile phone apps:** Given that cellular phones are now a critical medium of communication and work for professionals, we collected information from phones related to screen time, apps used and duration of usage, etc.

**Relating collected data based on specific organizing criteria**

Finally, we evolved a system for *relating* the data was collected on multiple criteria to put them into proper context. These organizing criteria fell into the following:

- **Time based:** Data is related based on time of day, hour, day, week, month, year, etc.
- **Organized groups:** Data is aggregated and related based on groups of subjects, by role (i.e., engineers versus managers), by geographic proximity, by organization, etc.

The diagram below summarizes the data collection process:



Depiction of process for gathering data from subjects

**Representing aspects of a scoring system: attributes, sub-attributes**

In order to properly introduce our scoring system, it is necessary to explain how factors that go into scoring were identified and computed.

To keep things simple, our scoring system envisioned having *top-level* scoring parameters called *attributes;* these attributes would roughly correspond to *categories* of work/apps that our subjects used in their day to day work activities (i.e. Emails, Service Desk, Business Apps Usage, etc.). The ultimate goal of the scoring system is to generate a number between 1 to 10 for each attribute; 10 indicating highest productivity being reached for that attribute, and 1 indicating a total failure to achieve productivity in that attribute.

For each attribute identified, specific *measurable criteria* (which may in fact apply to one or more attributes) are identified collectively as "fulfilling" the evaluations necessary for each attribute.

Taking the example of Business Apps Usage. It can be stated that scoring of "Business Apps Usage" as an attribute is a combination of:

- How much time was spent in a day working with apps
- What percentage of that time was spent on *actual business apps* (as opposed to "non-business" apps like games. Note here that "non-business" as an adjective is subjective, i.e. the professional may be a professional video games QA tester, in which case this adjective would not apply).

In the above example, two *sub-attributes* have been identified:

- Time spent on apps in a day
- Percentage of time spent in business apps versus non-business apps

**Introducing formulae and score tables for scoring**

The next step is to *map* possible ranges of values for each sub-attribute to entries in a *score table*. Take the example of the "percentage of time spent in business apps versus non-business apps" sub-attribute. A possible score table could look like this…

| **Percentage of time spent in business apps** | **Sub-attribute score** |
|---|---|
| 100% | 10 |
| 90% | 9 |
| 80% | 9 |
| 75% | 7 |
| 60% | 6 |
| 55% | 6 |
| 50% | 5 |

This can be interpreted in the following ways:

1. A subject who managed to spend 100% of his time exclusively working on business apps is a stellar achiever and gets the highest score (10).
2. A subject who spends between 80-90% of his time exclusively on business apps is pretty close to accomplishing the desired goal, he gets a score of 9.
3. When encountering subjects who have spent less than 80% of their time, the exact percentage determines how much their score is reduced (i.e. someone who doesn't spend more than half his working time using business apps gets a score of 5).

In addition, there may be a possibility that (for an attribute), one contributing sub-attribute may be *more important* than another contributing sub-attribute for a subject, or profile, or any other organizing criteria. For instance, it may be that the percentage of time spent on business apps (sub-attribute 1) may be more important than the total time spent on apps (sub-attribute 2), even if total time spent is part of the computation.

In this case, a scoring *weight* (a value between 0 to 1) is assigned to sub-attributes, such that the total *weight* of all sub-attribute weights is 1.

Carrying forward our last example, it would then be possible, say, to assign a weight of 0.2 to "total time spent on apps", but a weight of 0.8 to "percentage of time spent on business apps" (notice that the sum of combined weights evaluates to 1, this is important for our scoring system).

In such a scenario, the *attribute score* would be:
Attribute score = (*weight of sub-attribute 1 * score of sub-attribute 1)* + (*weight of sub-attribute 2 * score of sub-attribute 2)* + …. + (*weight of sub-attribute n * score of sub-attribute n)*
Where *n* = total number of sub-attributes that are used to compute the score for the attribute.

**Introducing bias and relevance for overall scores**

Now that each category of productivity was representable via attribute scores, further work was done to account for these factors:

- It's possible that attributes don't have the same ***relevance*** across job profiles. For example, it may be the case that a sales professional doesn't really rely on a Code Quality category to understand his productivity (when it would be **very** relevant to a software engineer), but may find more relevance in a good Emails score or Meetings score (for instance, measuring number of emails resulting in confirmed sales).
- Even in relevant attributes, a particular individual may have a ***bias*** for specific attributes over others. Take, for example, a support engineer who considers having a good

Service Desk score to be more important that Business Apps Usage.

For this reason, when looking at the *overall* score based on all attributes, we provide a formula that incorporates individual attributes' bias and relevance as follows:
Overall score = sum(attribute score * bias * relevance) / sum (relevance of each attribute)
This gives us an overall score (between 1 to 10) by which a subject can get a calculated assessment of overall productivity for a particular day.

**Scoring accuracy evaluation via "declared" scores collection from subjects**

In order to verify that calculated scores (as per our devised scoring system) were accurate in measuring productivity, subjects in the associated group were asked to maintain a day-to-day diary of their activities, primarily "declaring" a score between 1 to 10 for each day. The purpose of this exercise was to provide a "reference" point against which to assess computed scores to determine degree of accuracy.

**Representing aspects of a scoring system: attributes, sub-attributes**

In order to properly introduce our scoring system, it is necessary to explain how factors that go into scoring were identified and computed.

To keep things simple, our scoring system envisioned having *top-level* scoring parameters called *attributes;* these attributes would roughly correspond to *categories* of work/apps that our subjects used in their day to day work activities (i.e. Emails, Service Desk, Business Apps Usage, etc.). The ultimate goal of the scoring system is to generate a number between 1 to 10 for each attribute; 10 indicating highest productivity being reached for that attribute, and 1 indicating a total failure to achieve productivity in that attribute.

For each attribute identified, specific *measurable criteria* (which may in fact apply to one or more attributes) are identified collectively as "fulfilling" the evaluations necessary for each attribute.

Taking the example of Business Apps Usage. It can be stated that scoring of "Business Apps Usage" as an attribute is a combination of:

- How much time was spent in a day working with apps
- What percentage of that time was spent on *actual business apps* (as opposed to "non-business" apps like games. Note here that "non-business" as an adjective is subjective, i.e. the professional may be a professional video games QA tester, in which case this adjective would not apply).

In the above example, two *sub-attributes* have been identified:

- Time spent on apps in a day
- Percentage of time spent in business apps versus non-business apps

**Introducing formulae and score tables for scoring**

The next step is to *map* possible ranges of values for each sub-attribute to entries in a *score table*. Take the example of the "percentage of time spent in business apps versus non-business apps" sub-attribute. A possible score table could look like this…

| **Percentage of time spent in business apps** | **Sub-attribute score** |
|---|---|
| 100% | 10 |
| 90% | 9 |
| 80% | 9 |
| 75% | 7 |
| 60% | 6 |
| 55% | 6 |
| 50% | 5 |

This can be interpreted in the following ways:

4. A subject who managed to spend 100% of his time exclusively working on business apps is a stellar achiever and gets the highest score (10).
5. A subject who spends between 80-90% of his time exclusively on business apps is pretty close to accomplishing the desired goal, he gets a score of 9.
6. When encountering subjects who have spent less than 80% of their time, the exact percentage determines how much their score is reduced (i.e. someone who doesn't spend more than half his working time using business apps gets a score of 5).

In addition, there may be a possibility that (for an attribute), one contributing sub-attribute may be *more important* than another contributing sub-attribute for a subject, or profile, or any other organizing criteria. For instance, it may be that the percentage of time spent on business apps (sub-attribute 1) may be more important than the total time spent on apps (sub-attribute 2), even if total time spent is part of the computation.
In this case, a scoring *weight* (a value between 0 to 1) is assigned to sub-attributes, such that the total *weight* of all sub-attribute weights is 1.
Carrying forward our last example, it would then be possible, say, to assign a weight of 0.2 to "total time

spent on apps", but a weight of 0.8 to "percentage of time spent on business apps" (notice that the sum of combined weights evaluates to 1, this is important for our scoring system).
In such a scenario, the *attribute score* would be:
Attribute score = (*weight of sub-attribute 1 * score of sub-attribute 1)* + (*weight of sub-attribute 2 * score of sub-attribute 2)* + …. + (*weight of sub-attribute n * score of sub-attribute n)*
Where *n* = total number of sub-attributes that are used to compute the score for the attribute.

**Introducing bias and relevance for overall scores**
Now that each category of productivity was representable via attribute scores, further work was done to account for these factors:

- It's possible that attributes don't have the same ***relevance*** across job profiles. For example, it may be the case that a sales professional doesn't really rely on a Code Quality category to understand his productivity (when it would be **very** relevant to a software engineer), but may find more relevance in a good Emails score or Meetings score (for instance, measuring number of emails resulting in confirmed sales).
- Even in relevant attributes, a particular individual may have a ***bias*** for specific attributes over others. Take, for example, a support engineer who considers having a good Service Desk score to be more important that Business Apps Usage.

For this reason, when looking at the *overall* score based on all attributes, we provide a formula that incorporates individual attributes' bias and relevance as follows:
Overall score = sum(attribute score * bias * relevance) / sum (relevance of each attribute)
This gives us an overall score (between 1 to 10) by which a subject can get a calculated assessment of overall productivity for a particular day.

**Introducing concept of "automated" score improvement notifications**

In addition to the scoring system, the authors devised a mechanism for achieving the following:

- Continuous generation of scores based on "current" data at specific times of day. This would continue till the end of the day, at which point the "final" score for the day was computed
- When scores dipped below an optimal range, the system would randomly choose one of the points in the day to "prompt" the subject to undertake actions to improve the score (usually why an automated email)
- The subjects were given no specific instructions on whether to respond to the notification or not.

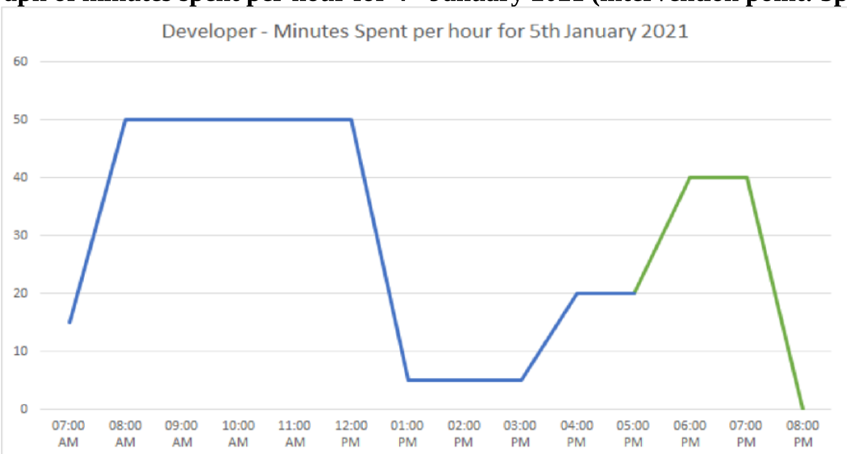
Assessment of changes in engagement post notifications via time intervention analysis. In order to determine whether there was a significant change (for the better) in engagement, the following methodology was utilized for collected data:

- Collected data was plotted in suitable intervals in a time series. Bucket size would vary depending on nature of data (i.e. hourly data for business apps usage, daily data for closed defects, etc.)
- Times of day at which automated notifications from the system "kicked in" were considered as *intervention points* in the time series
- A computation of the "average" value pre and post intervention was calculated and then analyzed for deviations.

The expectation was that such deviations would be found, and would be assessed to determine in what way the interventions had changed engagement outcomes.

## TABLES AND GRAPHS

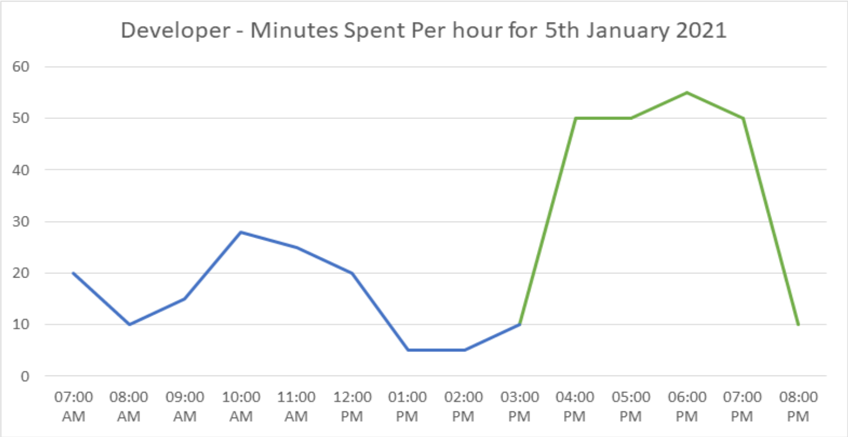
**1.1– Developer – Graph of minutes spent per hour for 4th January 2021 (intervention point: 5pm)**




For chart above, average "minutes per hour utilization" recorded for each phase:

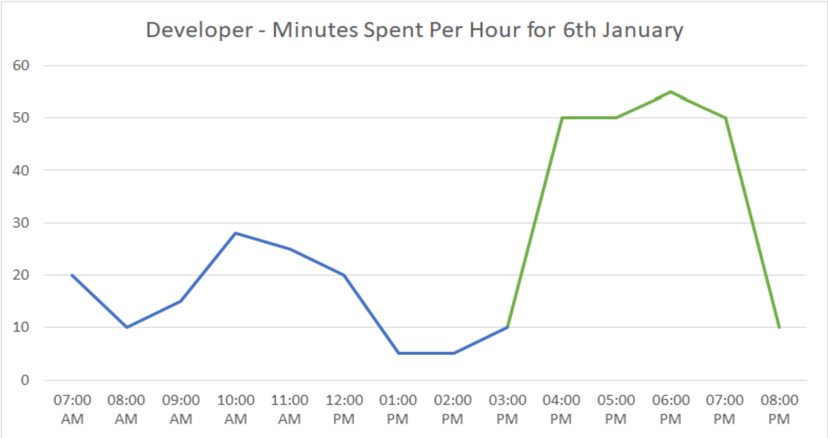
Prior to intervention: 15 minutes per hour utilization
Post intervention: 43.75 minutes per hour utilization (increase in average minutes per hour utilization).

**1.2 Developer – Graph of minutes spent per hour for 5th January 2021 (intervention point: 3pm)**



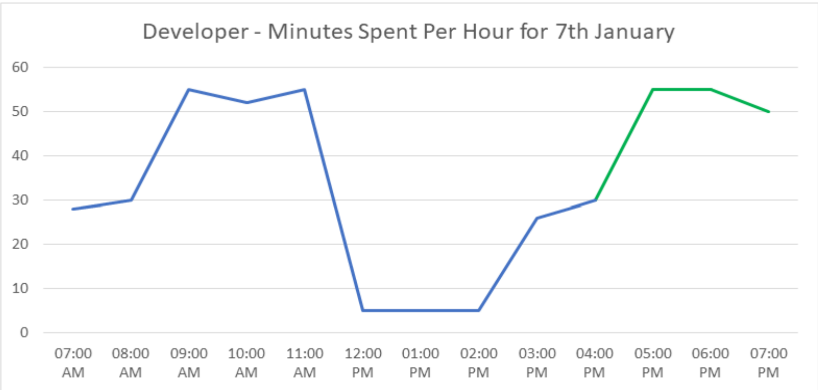

For chart above, average "minutes per hour utilization" recorded for each phase:
Prior to intervention: 29.09091 minutes per hour utilization
Post intervention: 40 minutes per hour utilization (increase in average minutes per hour utilization).

**1.3 Developer – Graph of minutes spent per hour for 6th January 2021 (intervention point: 3pm)**



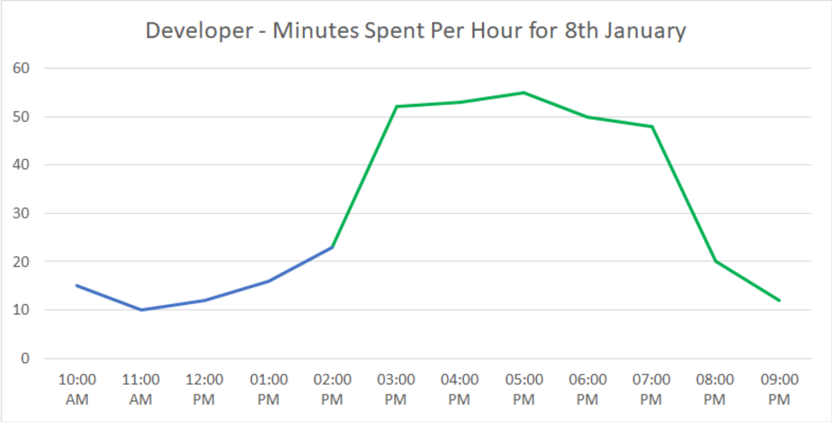

For chart above, average "minutes per hour utilization" recorded for each phase:
Prior to intervention: 16 minutes per hour utilization
Post intervention: 37.5 minutes per hour utilization (increase in average minutes per hour utilization).

**1.4 Developer – Graph of minutes spent per hour for 7th January 2021 (intervention point: 4pm)**



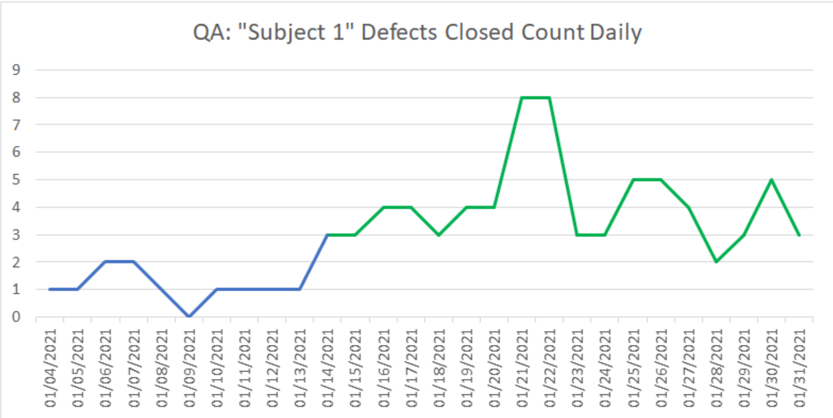

For chart above, average "minutes per hour utilization" recorded for each phase:
Prior to intervention: 29 minutes per hour utilization
Post intervention: 47.5 minutes per hour utilization (increase in average minutes per hour utilization).

**1.5 Developer – Graph of minutes spent per hour for 8th January 2021 (intervention point: 2pm)**



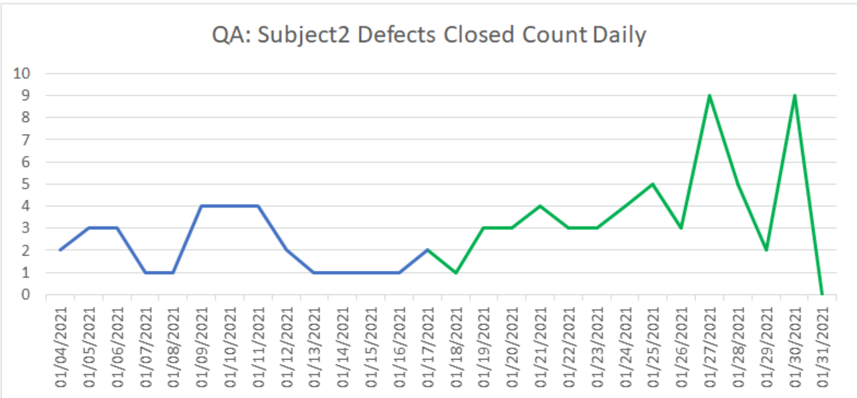

For chart above, average “minutes per hour utilization” recorded for each phase:
Prior to intervention: 13.25 minutes per hour utilization
Post intervention: 39.125 minutes per hour utilization (increase in average minutes per hour utilization).

**2.1 QA – “Subject 1” Defects Closed Count Daily from 4th January 2021 to 31st January 2021 (Intervention on 13th January 2021)**



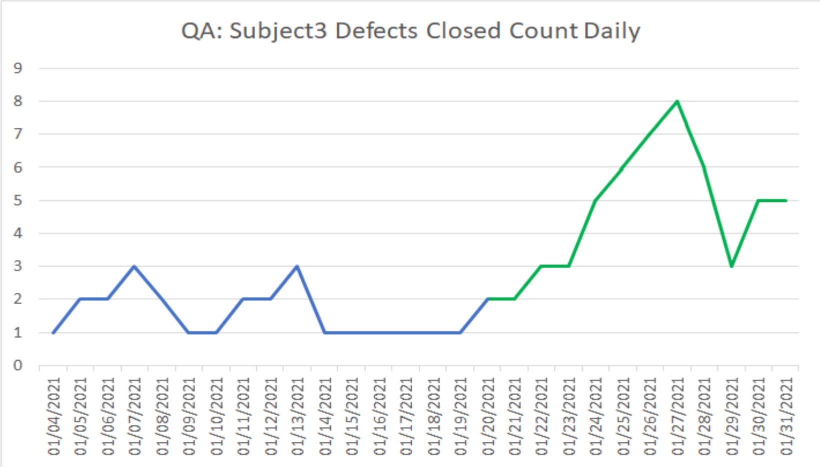

For chart above, average “defects closed count daily” recorded for each phase:
Prior to intervention: 1.11111111 defects closed count daily
Post intervention: 3.9473684 defects closed count daily (increase in defects closed count daily)

**2.2 QA – “Subject 2” Defects Closed Count Daily from 4th January 2021 to 31st January 2021 (Intervention on 18th January 2021)**



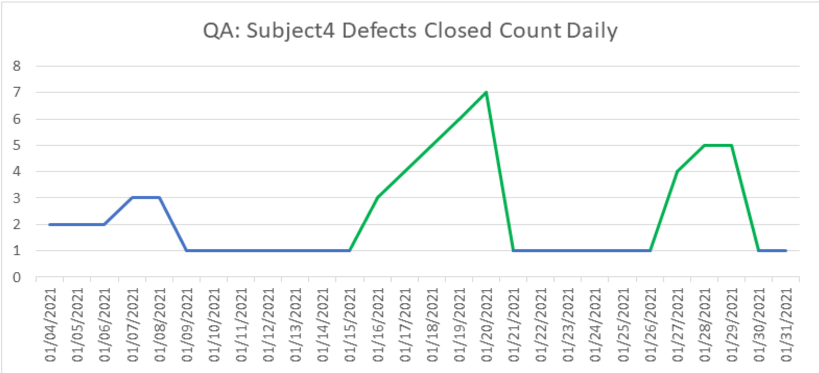

For chart above, average “defects closed count daily” recorded for each phase:
Prior to intervention: 2.14285714 defects closed count daily
Post intervention: 4.1538462 defects closed count daily (increase in defects closed count daily)

**2.3 QA – "Subject 3" Defects Closed Count Daily from 4th January 2021 to 31st January 2021 (Intervention on 20th January 2021)**



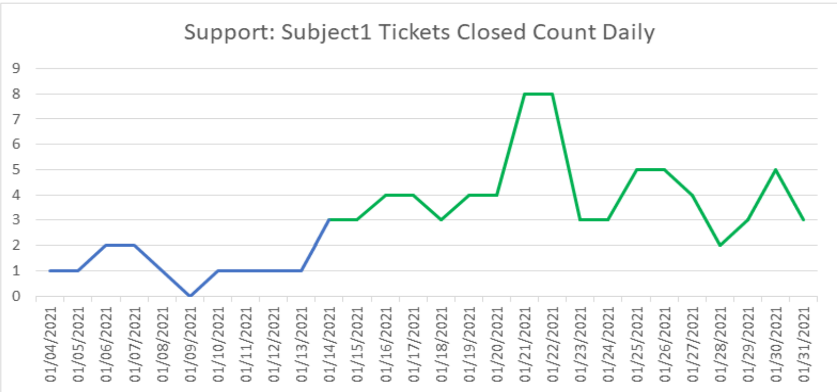

For chart above, average "defects closed count daily" recorded for each phase:
Prior to intervention: 1.5625 defects closed count daily
Post intervention: 4.5833333 defects closed count daily (increase in defects closed count daily)

**2.4 QA – "Subject 4" Defects Closed Count Daily from 4th January 2021 to 31st January 2021 (Interventions on 15th January 2021 and 26th January 2021)**



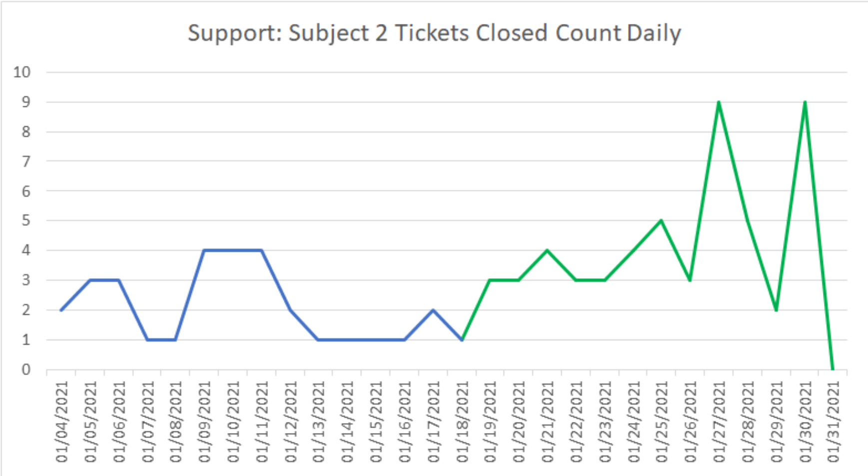

For chart above, average "defects closed count daily" recorded for each phase:
Prior to intervention: 1.63636364 defects closed count daily
Post intervention 1: 2.8181818 defects closed count daily (increase in defects closed count daily)
Post intervention 2: 2.83333333 defects closed count daily (increase in defects closed count daily)

**3.1 Support – "Subject 1" Tickets Closed Count Daily from 4th January 2021 to 31st January 2021 (Intervention on 13th January 2021)**



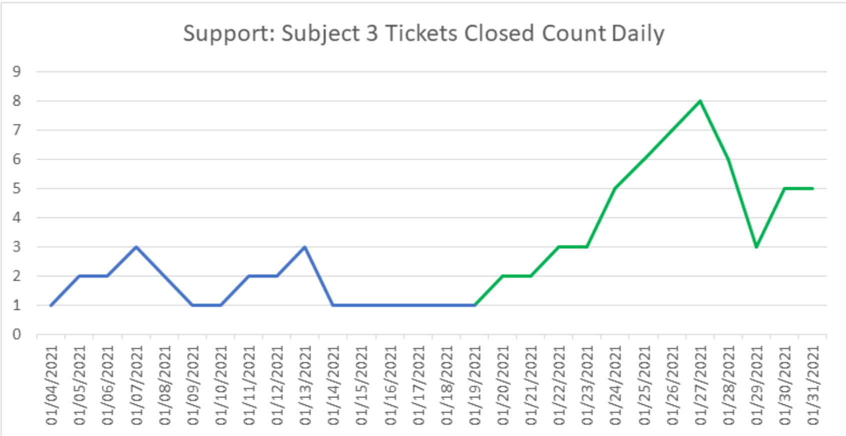

For chart above, average "tickets closed count daily" recorded for each phase:
Prior to intervention: 1.1111111 tickets closed count daily
Post intervention: 3.94736842 tickets closed count daily (**increase** in tickets closed count daily)

**3.2 Support – "Subject 2" Tickets Closed Count Daily from 4th January 2021 to 31st January 2021 (Intervention on 18th January 2021)**




For chart above, average "tickets closed count daily" recorded for each phase:
Prior to intervention: 2.1428571 tickets closed count daily
Post intervention: 3.85714286 tickets closed count daily (increase in tickets closed count daily)

**3.3 Support – "Subject 3" Tickets Closed Count Daily from 4th January 2021 to 31st January 2021 (Intervention on 20th January 2021)**




For chart above, average "tickets closed count daily" recorded for each phase:
Prior to intervention: 1.5625 tickets closed count daily
Post intervention: 4.58333333 tickets closed count daily (increase in tickets closed count daily)

## RESULTS

Statistically significant increase in engagement in group utilizing scoring system due to gamification.

For surveyed subjects, it was observed that the magnitude of engagement (i.e. increase in intensity, count, or any metric in general associated with activities relevant to job profile of surveyed individual) would increase on average after intervention by the scoring system compared to the magnitude of engagement prior to the intervention (though the actual magnitude would vary, the deviation in increase of magnitude can be adjusted for by factors including motivation levels of the subject, time available, etc.).

## CONCLUSIONS

Based on the results gathered from this study, the authors concluded that the alternate hypothesis had been proven (put plainly, a scoring system that results in interventions/notifications to individuals for the purpose of improving productivity is more likely to increase engagement from individuals). While the extent of the engagement would vary from individual to individual (and other factors like situation, time of day, nature of profile and activities, etc.) in general terms, it appeared that a significant accelerator to improving productivity/wellness would be the introduction of a scoring system that accurately assesses current productivity/wellness based on individual's data, and provides clear and concrete notifications tailored towards activities that contribute towards productivity/wellness. By removing the uncertainty and anxiety that comes from lack of self-efficacy, unclear goals and uncertain progress, organizations who wish to encourage a culture of sustainable productivity combined with focussed wellness initiatives would do well to pay attention to such systems.

## REFERENCES


- Saad, D. (2018). Impact of employee motivation on work performance. *International Journal of Scientific and Research Publications*, *8*(3), 295-308.
- Bagul, D. B. (2014). Study of employee's performance appraisal system. *Scholarly Research Journal of Humanity Sciences and Languages*.
- Khajuria, G., & Khan, N. (2022). Perceived Organisational Support and Employee Engagement: A Literature Review. *Journal of Positive School Psychology*, *6*(3), 1366-1384.
- Knox, S., & Marin-Cadavid, C. (2022). A practice approach to fostering employee engagement in innovation initiatives in public service organisations. *Public Management Review*, 1-26.
- Han, X., Chen, S., & Chen, B. (2022). From employee engagement to customer engagement: A multilevel dual-path model of engagement spillover effects in service ecosystems. *Journal of Retailing and Consumer Services*, *64*, 102815.
- Antony, M. R. (2018). Paradigm shift in employee engagement–A critical analysis on the drivers of employee engagement. *International Journal of Information, Business and Management*, *10*(2), 32-46.
- Bedarkar, M., & Pandita, D. (2014). A study on the drivers of employee engagement impacting employee performance. *Procedia-Social and Behavioral Sciences*, *133*, 106-115.
- Tomlinson, G. (2010). Building a culture of high employee engagement. *Strategic HR review*.
- Shaik, F. F., & Makhecha, U. P. (2019). Drivers of employee engagement in global virtual teams. *Australasian Journal of Information Systems*, *23*.
- Crabb, S. (2011). The use of coaching principles to foster employee engagement. *The Coaching Psychologist*, *7*(1), 27-34.


**Cite This Article:** Nitish Shrivastava, Neel Tiwari, Elaiyaraja Thangavel (2022). A Comprehensive Scoring System for Productivity Increases Engagement and Wellness Outcomes. *East African Scholars J Eng Comput Sci*, *5*(5), 55-68.