Challenges to consider in monitoring litter prevention

Prepared by: Daniel Stunell and David Barnes

Date: June 2017

European Union
gov.scot

EUROPE & SCOTLAND
European Regional Development Fund
Investing in a Smart, Sustainable and Inclusive Future

Contents

Introduction

Litter has significant negative social, environmental and economic impacts. Preventing litter from occurring is the most efficient and effective way of minimising this harm and maximising the value from these wasted resources[1].

One of the main challenges faced in preventing litter is trying to identify what interventions are successful in achieving the necessary long term behaviour change. This difficulty is a result of:

- The fact litter consists of very many small items, which are widely but not evenly deposited, and may also be subject to significant natural variation over time;
- The complex combination of factors driving behaviour[2] may mean that even successful interventions may not make a sufficient difference to be robustly evidenced via some techniques.

Nonetheless, monitoring is a highly desirable component of any anti-litter intervention. However, there is no one-size-fits all.

A key test you should apply to monitoring is **proportionality**. If you are transferring a "tried-and-tested" approach from a comparable context somewhere else, then you will probably expect it to work. In this case, the main purpose of monitoring would simply be to double check performance was as expected, and a minimalist approach might be appropriate.

On the other hand, if you are trying a new approach, or working in a new context, then extensive monitoring may be in order, as you need to know if the idea is working or not, and be able to take steps to improve or optimise it. Similarly, interventions with a high degree of money or time invested in them are likely to justify a higher level of monitoring (as there is more at stake).

You also need to be aware of **confirmation bias** when assessing activities. It is human nature not to question results if you get the outcome you want but look robustly at external factors if you don't get the results you were looking for. As an example you should be considering the impact of weather on an intervention regardless of whether litter has increased or decreased.

There are often a **chain of impacts** that ultimately lead to litter prevention. Often the earlier links in the chain will be the easiest to measure – but it will be harder to conclude the intervention is having the ultimate end effect from just these early links in the chain. On the other hand, the final stage in the chain is very likely to be influenced by external factors, making conclusions difficult to draw too. Measuring multiple indicators along the chain will often be expensive.

The **purpose of this paper** is not to provide the necessary skills to undertake robust monitoring of litter interventions. Instead it is intended to highlight the areas that should be considered, to allow readers to ask informed questions of interventions that are presented to them and encourage either the development of these skills or procuring of the necessary skills when designing new projects.

Zero Waste Scotland are interested in hearing about new approaches to monitoring and continuing to refine this paper. If you would like to contribute or provide feedback then please email litterandflytipping@zerowastescotland.org.uk

Consideration 1: Ensure the right thing is being monitored

How is your intervention is expected to make a difference? For example a communications campaign might be expected to raise awareness, meaning the public drop less litter, meaning there is less litter on the ground. Awareness, the act of dropping litter, or litter on the ground could all be measured.

[1] Scotland's Litter Problem, Quantifying the scale and cost of litter and flytipping
[2] Individual, Social and Material – Influencing Behaviours

Example: Improving the deterrent effect of enforcement by increasing the payment rate for Fixed Penalty Notices (FPN). In this case payment rates would be an appropriate measure of success (and easier to measure than an abstract concept like “deterrence”). In contrast, litter levels might be quite far removed from your intervention, and thus less useful (and more expensive) to measure.

Consideration 2: Short term vs Long Term impacts

How long will change last? If improvements last for only a short time, then cost-effectiveness will be much reduced. So you may want to understand how long improvements endure for, both to help select interventions for the future, and to decide if and when to repeat an intervention. Be especially aware of interventions which might be expected to have a natural lifespan.

Example: A school intervention with final year students might have limited effect once those students leave and new ones arrive unless the intervention is embedded or passed on within the school as an institution. Equally, you may want to leverage significant short term impacts to underpin longer lasting communications campaigns e.g. as a PR focused activity.

Consideration 3: Normal variability over time

How much fluctuation is there in what you are measuring? Almost all activity or behaviour will have some variation over time and be influenced by a variety of factors e.g. footfall or resources. The greater natural variation is in a specific context, the harder it will be to reliably measure change as random changes will drown out small long term changes unless you gain a lot of data over long time periods.

Example: Where interventions are targeted to influencing behaviour in a specific location then it can make regular and accurate measuring of litter on the ground feasible. It is important however to consider other factors that could influence whether this increases or decreases at these monitoring points e.g. at a tourist attraction the weather could influence the number of visitors.

Try to avoid this by understanding your context in the first place and designing your monitoring to reduce this risk (for example: if weather is likely to be a key variable, only measure on comparable days; be aware of match days near football grounds; etc.).

Consideration 4: Sample selection

This links closely to consideration 1. What you *measure* needs to be representative of what you are trying to *influence*. This might be by geography (e.g. residential/commercial areas, SIMD), socio-demographics (e.g. age, or employment status). There are a range of standard methods for randomly sampling for surveys of people; approaches to cleanliness surveys are more bespoke. But in all cases, you can only generalise from what you have measured to the bigger picture if you are confident the results you have will be typical of both.

Example: If your intervention targets a whole community’s awareness of litter, you need to make sure the people you ask about their awareness are typical for the whole community. Asking only schoolchildren, or only people on the high street, won’t do this. (But a high street survey might be a good way to understand how *users of the high street* see things).

Consideration 5: Eliminating other influences

Is it possible other factors influenced what you are measuring? You aren’t in a laboratory and the intervention will be subject to other factors that you have no control over (and may not even be aware of). Recording complicating factors (such as weather, time of day, day of the week, school holidays) can help you to spot this. Designing your monitoring to minimise this can be invaluable (always measuring at the same time of day, or just before street cleansing occurs for example). If you have the option to create a control group or control area (where you do not intervene) this can be a useful

contrast to where you do – if litter drops in both areas, this probably isn't down to your intervention. But control areas are hard to select and manage – are they really the same? Is it really independent of any intervention effects?

Example: Where an intervention is being run in a secondary school then a control group could be established my monitoring a similar school, with agreed characteristics, without the intervention taking place. This might provide insight into natural variability of levels of litter, impacts of weather etc. But both your intervention and your control are still subject to random events.

Control groups work best with larger numbers – for example if you issue 1000 FPN reminder letters and change the wording in half of them before randomly sending them out, you are likely to have eliminated external influences, which should cancel out. Differences between the groups should then be real.

Consideration 6: Scalability / Sustainability

Can the intervention be rolled out with existing resources or has the source of additional resources been considered? Significant impacts can be delivered over the short term or in a limited area by increasing the resources used. This may not be possible for an entire area and so it is important to consider how interventions could be "scaled up" or focus on short-term benefits e.g. increased PR.

Example: Increasing the presence of enforcement staff in response to a specific issue is a common occurrence but it often isn't a long term solution as it requires reallocating resources from elsewhere. It is a powerful tool to raise awareness of the potential consequences of littering and can amplify the impact of further actions e.g. a follow on communications campaign.

Consideration 7: Transferability

Can the intervention be transferred to different locations? Success in one type of location does not automatically mean that it will work or have the sample impact elsewhere. Consider what is likely to influence the success of the intervention and whether those exist where it is being proposed. Be alert to unique success factors – community interventions often rely on a key individual for contacts, to motivate others, or simply for sheer commitment. This isn't always transferable or replicable elsewhere.

Example: Zero Waste Scotland research[3] has found that littering behaviour is context specific and therefore the messages to address this effectively should be too. This means that messages targeting residential areas are likely to work in any residential area across the country but not in city centre locations.

Consideration 8: Quantity of data

Is there enough information to make conclusions with confidence? More data is a way to overcome some of the challenges highlighted above – longer baselines or follow up periods can reduce the impact of natural variation, which averages out over time. Bigger samples will almost always be better to generalise from – a survey of 1000 people can be used to estimate results for all of Scotland, provided they are representative. A sample of 10 can't, no matter what you do.

But don't just assume more is better – quality is essential too. If there is another design flaw in your monitoring (such as an unrepresentative sample), more data won't help.

Example: Understanding the items that make up the litter waste stream is extremely useful information in helping to take preventative action i.e. if you know what is there then can start to look at potential sources and work to address these. But it is important not to draw conclusions about a wider area based on a small number of samples.

[3] Public Perceptions and concerns around litter, Zero Waste Scotland

Consideration 9: The Eighty: Twenty Rule

The Pareto principle isn't really a rule, but it's a useful rule of thumb in many contexts, and particularly helpful in terms of proportionality of monitoring. This states that 80% of the effects come from 20% of the causes, so identifying this 20% in your intervention design is vital. This will allow you to ensure that the monitoring requirement is targeted, to capture the majority of the effects from the intervention, while still being proportionate in terms of cost, resources etc.

Example: In assessing the impact of the impacts of the carrier bag charge, good data was available both pre and post charge for the 7 major grocery retailers. Although a very small percentage of overall retail outlets in Scotland they were responsible for the majority of single use carrier bags distributed. They are therefore a very useful indicator of the overall impact of this regulation.

Consideration 10: Measurement sensitivity

Are your expectations of what the intervention will achieve realistic? And have you matched your choice of what/how you are measuring to this expectation? This also brings us back to consideration 1 – you need to choose the right thing to measure. Interventions with a bigger impact should be easier to measure as their true effects are less likely to be hidden by random variations.

Example: A community engagement that reaches 200 people in a couple of residential streets around a communal green space might make a significant localised difference in a subsequent litter count or cleanliness assessment. 200 people engaged in the same way but who live scattered throughout a larger town may not make a measurable difference at either a localised or town level, even if the net effect of the interventions in terms of items not littered is the same. Only the former is likely to be *intensive* enough to make a *measurable* difference.

Consideration 11: Logic testing

Are you missing any links in the chain of how your intervention will make a difference? Testing this is simply about looking self-critically at your intervention design, not about doing extra measurement, but it can be easy to forget once you get focused on the detail of intervention design. Writing out the "links in the chain" that need to happen for your activity to have maximum effect can be extremely helpful.

Example: You increase enforcement activity to deter litterers, and you measure increased enforcement activity. This is a sensible measure of success, as far as it goes. But how will this additional activity deter others? Do you have communications or other awareness activity in place to ensure they get the message that litterers are more likely to be caught and punished? You may choose to measure the communication activity too – but if this activity is missing altogether, you can already guess your intervention will be less effective.

Conclusion

Monitoring litter prevention is challenging and no one size fits all. However by investing the time and resources to implement a robust monitoring and evaluation strategy, when designing projects and/or assessing information being presented to you, stakeholders can make more informed decisions. This will result in adopted proposals targeting opportunities that are important for your specific circumstances and which robustly demonstrate positive impacts.

Monitoring performance has also been demonstrated to improve performance, by reinforcing what is working and identifying areas for optimisation. It will then be a collection of these interventions that result in a significant reduction, and eventual elimination, of littering behaviour.

01786 433930 zerowastescotland.org.uk @ZeroWasteScot