Social Science & Medicine 181 (2017) 54–65

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Changing health care provider performance through measurement

Kenneth L. Leonard [a, *], Melkiory C. Masatu [b]

[a] Department of Agricultural and Resource Economics, University of Maryland, College Park, MD 20740, USA
[b] Centre for Educational Health in Development, Arusha (CEDHA), PO Box 1162, Arusha, Tanzania

ARTICLE INFO

Article history:
Received 19 March 2016
Received in revised form
6 March 2017
Accepted 18 March 2017
Available online 22 March 2017

Keywords:
Quality of care
Low income countries
Results based financing
The Hawthorne effect
Altruism
Pro-social motivation
Supportive supervision

ABSTRACT

Can the quality of care be improved by repeated measurement? We show that measuring protocol adherence repeatedly over ten weeks leads to significant improvements in quality immediately and up to 18 months later without any additional training, equipment, supplies or material incentives. 96 clinicians took part in a study which included information, encouragement, scrutiny and repeated contact with the research team measuring quality. We examine protocol adherence over the course of the study and for 45 of the original clinicians 18 months after the conclusion of the project. Health workers change their behavior significantly over the course of the study, and even eighteen months later demonstrate a five percentage point improvement in quality. The dynamics of clinicians' reactions to this intervention suggest that quality can be improved by the repeated measurement by external peers in a way that provides reminders of expectations.

1. Introduction

In developing and transition countries, quality of health care services are generally low in large part because adherence to medical protocols is low (Das and Gertler, 2007; Das and Hammer, 2007; Holloway et al., 2013; Leonard and Masatu, 2007; Rowe et al., 2005) and increases in adherence are one of the most effective ways to improve outcomes and prevent childhood deaths (Black et al., 2003; Jones et al., 2003; Rowe et al., 2005). Many types of interventions to improve quality have been tested in developing countries with somewhat positive results, including improved supervision, additional training, and interventions to change the workplace culture (such as institutional and management changes and group-based techniques). Yet few of these techniques have lasting impacts or have been rolled out effectively at scale and low adherence remains. Importantly there is increasing evidence of an important know-do gap (Das and Hammer, 2007; Leonard and Masatu, 2010; Rethans et al., 1991) in which clinicians demonstrate that they know how to adhere to protocol but choose not to do so. Indeed it is possible that the short-term benefits to a wide variety of interventions and the demonstrated know-do gap are connected and demonstrate a basic Hawthorne effect: when faced with the immediate attention and scrutiny inherent in *any* intervention, clinicians improve adherence but adherence falls as the attention diminishes.

This paper examines a specific program designed to extend a short-term Hawthorne effect to the medium-term by maintaining scrutiny and attention without any training, explicit supervision, institutional reforms or external rewards. The study encouraged clinicians to adhere to protocol and then returned about two weeks later to see if quality had improved. To test for the medium-term impact we designed the intervention with a follow-up at about 6 weeks after the original encouragement. As we show in this paper, we observed the opposite of what we expected: quality was only marginally higher in the short-term (the two week window) but significantly higher for the medium-term (the 8 week window). As a result of this surprising finding, we returned almost a year and half later to visit the same clinicians and found that quality was still higher than in the baseline though slightly less than at the medium-term.

This paper uses this data on the response to the program in the short, medium and long-term and shows evidence that the know-do gap is closed with increased and sustained attention and scrutiny, and that some forms of measurement, by themselves, can lead to long-term improvements in quality.

* Corresponding author.
E-mail address: kleonard@arec.umd.edu (K.L. Leonard).

http://dx.doi.org/10.1016/j.socscimed.2017.03.041
0277-9536/

2. Methods

The data in this paper comes from 4512 patient exit interviews conducted in two different periods and over three samples of clinicians. For a four-hour window on a randomly-selected unannounced date, members of the enumeration team asked all the patients who had visited a particular clinician a series of questions about their consultation based on the symptoms that they reported. The interviews with patients followed the Retrospective Consultation Review (RCR) instrument which allows us to reconstruct the clinicians' activities, specifically the extent to which they followed protocol (Brock et al., 2016). All the questions used in the instrument are reported in Table 6, Table 7, and Table 8 in the appendix. For example, according to protocol, a doctor is supposed to check for neck stiffness for all patients reporting with a fever, and we ask patients with a fever "did the doctor ask you if your neck was stiff?" Although patient recall is not perfect, it is highly correlated with what actually takes place (Leonard and Masatu, 2006).

The first period of data collection represents the initial study which took place over approximately 10 weeks (for each health worker). The second period, a follow-up study, took place over a year later. In the Period 1, we visited all clinics in the urban and peri-urban areas of Arusha, Tanzania and enrolled all clinicians that we could find working in the OPD clinics of facilities that had a reasonable number of patients per day (at least 5). This resulted in a sample of 96 clinicians in 40 health care facilities including clinicians working in public, private, and non-profit/charitable facilities. The term "clinician" refers to primary health workers who provide outpatient care. All clinicians have significant medical training but the majority of them do not have full medical degrees.

All clinicians in the sample consented to be in a study (103 clinicians were contacted but 1 refused consent and 5 consented but

Table 1
Sample averages at baseline.

	Samples					
	baseline	treatment	comparison	p-value of comparison		
	(1)	(2)	(3)	1 vs. 2	1 vs. 3	2 vs. 3
Characteristics of the average health worker						
Individual Health Worker Protocol Adherence Scores in the baseline						
Overall	0.755	0.740	0.732	0.667	0.394	0.818
	(0.019)	(0.031)	(0.019)			
Primed items	0.741	0.709	0.715	0.450	0.396	0.879
	(0.023)	(0.036)	(0.020)			
Non-primed Items	0.780	0.768	0.799	0.642	0.358	0.224
	(0.015)	(0.023)	(0.014)			
Overall effect size	0.698	0.676	0.664	0.567	0.245	0.747
	(0.021)	(0.034)	(0.020)			
Observations	96	45	97			
Characteristics of the average consultation						
Baseline Protocol Adherence for consultations seen at baseline and after treatment						
Overall	0.728	0.729	0.730	0.968	0.879	0.937
	(0.008)	(0.012)	(0.007)			
Primed items	0.712	0.703	0.706	0.584	0.595	0.852
	(0.009)	(0.014)	(0.007)			
Non-primed Items	0.759	0.761	0.799	0.829	**0.000**	**0.000**
	(0.006)	(0.009)	(0.005)			
Overall effect size	0.666	0.667	0.662	0.975	0.729	0.754
	(0.009)	(0.013)	(0.008)			
Characteristics of the patient and caregiver						
gender (female = 1)	0.534	0.573	0.543	0.276	0.762	0.402
	(0.021)	(0.029)	(0.021)			
Age	21.734	23.098	23.586	0.217	**0.046**	0.664
	(0.648)	(0.890)	(0.664)			
Child (5–15)	0.181	0.187	0.162	0.840	0.389	0.357
	(0.016)	(0.022)	(0.015)			
Adult (15+)	0.639	0.672	0.680	0.325	0.135	0.804
	(0.020)	(0.027)	(0.019)			
Age of caretaker	31.230	32.505	32.241	**0.068**	**0.084**	0.709
	(0.410)	(0.561)	(0.417)			
Gender of caretaker	0.654	0.693	0.658	0.243	0.887	0.294
(female = 1)	(0.020)	(0.027)	(0.020)			
Fever	0.259	0.226	0.222	0.278	0.137	0.892
	(0.018)	(0.024)	(0.017)			
Cough	0.178	0.187	0.154	0.743	0.267	0.208
	(0.016)	(0.022)	(0.015)			
Diarrhea	0.059	0.069	0.063	0.577	0.779	0.748
	(0.010)	(0.015)	(0.010)			
Infant fever	0.100	0.108	0.091	0.702	0.584	0.400
	(0.012)	(0.018)	(0.012)			
Infant cough	0.073	0.075	0.063	0.891	0.513	0.493
	(0.011)	(0.015)	(0.010)			
Infant diarrhea	0.019	0.039	0.034	**0.064**	**0.097**	0.695
	(0.006)	(0.011)	(0.008)			
time of day	11.220	11.316	11.373	0.365	**0.094**	0.605
	(0.063)	(0.084)	(0.066)			
Observations	590	305	585			