TV Channels in Your Pocket! Linking Smart Pockets to Smart TVs

Figure 1: Smart-Pockets for television enable quick and convenient access to one's preferred TV content by implementing the "pocket metaphor": some TV channels are linked to specific pockets on the user's clothes, *e.g.*, a direct link to the Comedy Central channel can be stored in the trousers right front pocket, while a link to BBC News in the shirt breast pocket. In this figure, the user points his right arm to the trousers right front pocket and then towards the TV screen to start playing the TV show or channel previously linked to that pocket. The reverse operation is also possible: pointing to the TV and then back to a specific pocket will create a direct link between that pocket and the currently playing TV channel for fast later access.

Irina Popovici
MintViz Lab | MANSiD Research Center
University Ştefan cel Mare of Suceava
Suceava 720229, Romania
i.popovici.irina@gmail.com

Radu-Daniel Vatavu
MintViz Lab | MANSiD Research Center
University Ştefan cel Mare of Suceava
Suceava 720229, Romania
radu.vatavu@usm.ro

Wenjun Wu
State Key Lab of Software Development
Environment, Beihang University
Beijing, China
wwj@nlsde.buaa.edu.cn

ABSTRACT

We present a gesture-based user interface for smart TVs that employs deictic gestures to control the content displayed on the TV screen. Our interface implements an instance of the "Smart-Pockets" interaction technique, where links to digital content, in our case to users' preferred television channels and shows, are stored inside users' pockets and readily accessible with a mere pointing of the hand to those pockets. Pointing gestures to the pockets and towards the TV screen are detected using the Inertial Measurement Unit embedded in Myo, a smart armband. We discuss the ways in which our prototype opens new opportunities for hybrid, gesture- and pointing-based interactions for smart TVs as well as opportunities for designing interactions that take place at the periphery of user attention.

TVX '19, June 5–7, 2019, Salford (Manchester), United Kingdom

ACM ISBN 978-1-4503-6017-3/19/06.
https://doi.org/10.1145/3317697.3325119

BACKGROUND: SMART-POCKETS

"Smart-Pockets" represents a concept and interaction technique introduced by Vatavu [18] that employs links between digital content of any kind (*e.g.*, a list of emails, files, videos, etc., stored in the cloud) and physical locations on the user's body in the form of pockets. The concept was inspired by the fact that the most commonplace for storing one's physical belongings on the body is inside pockets, which have been devised over many years of fashion design to be both aesthetic and functional for a variety of contexts of use. However, **pockets could also be employed to store one's personal digital content**, such as a list of documents, phone contacts, or one's email inbox; see [18] for examples of applications. The digital content linked to pockets is displayed on an ambient screen by taking it out of the pockets and pointing the hand towards the screen.

Characteristics of Smart-Pockets gestures

Smart-Pockets gestures are fast, demand little to no attentional resources and precision to perform effectively, no training, and are easily extensible to other locations on the human body and to physical objects, such as "Smart-Containers;" see [18] for details.

Implementation: original and ours

In the original implementation [18], Smart-Pockets pointing actions were detected and recognized using the Microsoft Kinect sensor. In this work, we present a prototype that employs a wearable device, the Myo armband [10]. Our new implementation is highly portable and does not require any sensing infrastructure to be installed in the physical environment.

CCS CONCEPTS

• **Human-centered computing → Pointing**; **Gestural input**; **Ubiquitous computing**.

KEYWORDS

Smart TVs; Gesture input; Deictic gestures; Pointing; Smart-Pockets; Myo armband; Wearables; Peripheral interaction; Interactive TV.

ACM Reference Format:

Irina Popovici, Radu-Daniel Vatavu, and Wenjun Wu. 2019. TV Channels in Your Pocket! Linking Smart Pockets to Smart TVs. In *ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19), June 5–7, 2019, Salford (Manchester), United Kingdom.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3317697.3325119

INTRODUCTION

Current modalities to control the functions of the smart TV range from the simple action of pressing buttons on a remote control to using dedicated applications specifically designed for tablets and smartphones, and to gesture and voice input [1, 3, 4, 9, 16, 19]. The standard remote control, obsolete in many ways, has also been the subject of continuous research and development. For example, prior work proposed enhanced remote controls with embedded sensors, such as video cameras and accelerometers [1, 16, 17], or TV remote designs that sensed hand poses and mid-air input in the immediate space around them, such as in the Leap Motion gesture elicitation study of Vatavu and Zaiţi [19] conducted to inform design of low-effort gesture input for smart TV control.

In this paper, we propose a gesture-based interface for controlling the content displayed on the TV screen by relying on (1) the concept of "Smart-Pockets" [18] for linking digital content to specific areas on the user's body and (2) mid-air input in the form of deictic gestures (*i.e.*, pointing actions) recognized using a wearable device [11]. Our prototype, based on the Myo armband [10], enables viewers to quickly access their preferred TV channels, shows, or recorded TV content by taking that content out of their pockets and pointing it in the direction of the TV screen; see Figure 1 on the first page. Similarly, new associations between a TV channel and some pocket can be created by performing the reverse action: the hand pointed in the direction of the TV screen is put into a specific pocket (see Figure 2) to create a direct link between the TV channel and that pocket. We use our prototype to discuss implications of our work for a new kind of gesture-based user interface for the TV, little explored so far in the TVX community, that (1) connects digital TV content with physical, tangible objects and (2) creates the premises for designing interaction techniques at the periphery of user attention [2], suitable for low-effort input for TV control and lean-back television watching.

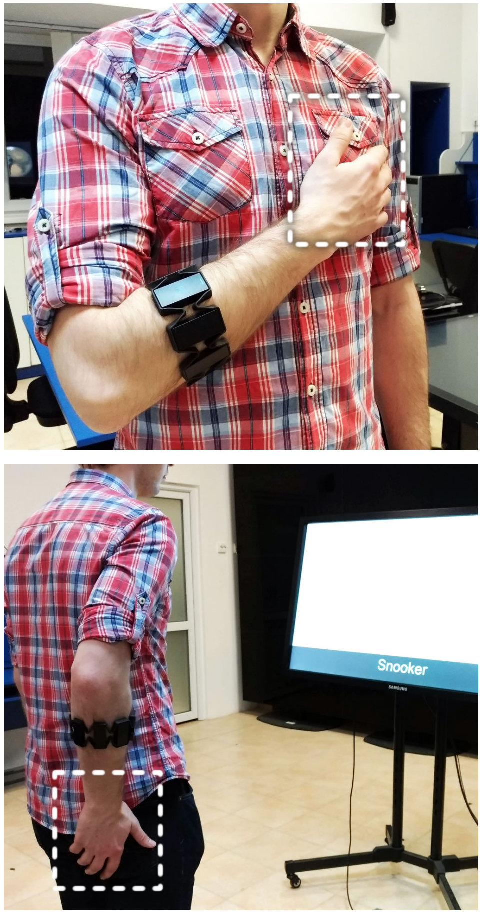

Figure 2: Two examples of pointing actions involving two "smart pockets": accessing the shirt left breast pocket (top figure) and the trousers rear right pocket (bottom figure). Note that pockets are not instrumented; instead, the smart armband detects the orientation of the hand and, thus, estimates the location of the pocket.

BACKGROUND: GESTURE-BASED INPUT FOR CONTROLLING THE TV

Previous work on interaction techniques for smart TVs focused on enhanced TV remote controls [1, 17], identifying suitable gesture commands for controlling the functions of the TV set [4, 16, 19], or on conducting studies to understand user performance, such as comparing the relative merits of various input modalities [3]. One direction contoured in the literature is represented by efforts to replace the standard TV remote control with other, "natural input" modalities, among which voice input [12] and gesture-based commands [3, 4, 19] represent the two most common options. For example, prior work has focused on designing suitable gesture commands for controlling the TV, to which end several gesture elicitation studies were conducted [4, 16, 19] to understand users' preferences for what types of gestures they would like to use to control their TVs. Another form of gesture input is represented by pointing. For example, the "Around-TV" system [17] used video projections to display multiple virtual screens around the TV set; the projected TV screens were manipulated in terms of their location and size on the wall behind the TV set using pointing actions performed using the Nintendo Wii remote control. Bobeth *et al.* [3] conducted a study to understand the differences in user performance and preference over three distinct input modalities for controlling the TV: an application on the tablet, free-hand gestures, and the standard TV remote control. Their study reported that the tablet condition (*i.e.*, tablet input mirrored on the TV screen) showed the best performance and was preferred by young adults and elderly people alike. One possible reason for this result is that free-hand and whole-body gesture input, as well voice input, come with several drawbacks, such as misrecognition errors or requirements for users to memorize commands when those commands are not intuitive. Therefore, there is still room to explore gesture-based interaction techniques for controlling the TV effectively with little cognitive effort and, ideally, at the periphery of user attention [2, 6]. In this work, we adopt the concept of "Smart-Pockets" [18] (see the side bar on the previous page for a detailed description) and present a wearable prototype that employs pointing actions and associations between preferred TV channels and various pockets on clothes.

PROTOTYPE

We present a prototype for controlling content on the TV screen by employing the Smart-Pockets [18] concept with a new implementation using a wearable device, the Myo armband [10]. Compared to the original implementation of Smart-Pockets (using the Microsoft Kinect sensor), our prototype is highly portable and does not require installation of any sensing infrastructure in the environment.

Apparatus and Development Tools

We developed a user interface for the TV using HTML 5, CSS 3 and JavaScript 1.7, and we tested it using the Google Chrome web browser (v72.0.3626.121) that ran on a laptop PC connected to a

THE MYO ARMBAND

Technical specifications

Myo weighs 93 grams, has 8 built-in EMG sensors, a 9-axis IMU, Bluetooth, delivers vibrotactile feedback, and is powered by an ARM Cortex M4 CPU; see https://support.getmyo.com/hc/en-us/articles/202648103-Myo-Gesture-Control-Armband-tech-specs

Applications

Prior work used Myo for various applications, such as in healthcare to collect patient data during training [8], video games [14] designed to encourage patients that need prosthetics to contract specific muscles, and for Virtual Reality [15]. Popovici and Vatavu [11] evaluated a gesture user interface for a TV menu with nine shortcuts to TV channels distributed in mid-air in front of the user's body. Their results showed good usability and high desirability, but high perceived workload.

Myo as a rapid prototyping device

Thalmic Labs (now North) ended production of Myo (https://support.getmyo.com/hc/en-us) in October 2018, but support is still available for practitioners and for the tens of thousands of devices out there (a figure estimated from https://mobilesyrup.com/2018/10/13/thalmic-labs-myo-ends-sales-next-project/). Myo remains a remarkable wearable device for capturing finger, hand, and arm gestures and especially for **rapid prototyping and evaluation** of new interface ideas and concepts, such as our Smart-Pockets implementation for TV control.

[1] https://support.getmyo.com/hc/en-us/articles/202647853-What-gestures-does-the-Myo-armband-recognize-

large, 55-inch Samsung UE55D smart TV; see Figure 1 on the first page. For demonstrative and testing purposes, we simulated television with digital content streamed from YouTube using HTML 5 and JavaScript API. The communication between the application running in the web browser and the Myo armband was implemented via Bluetooth using the Myo JavaScript SDK provided by North, former Thalmic Labs [10]. The Myo armband is a wearable device that captures electromyography (EMG) data at the forearm level, which is processed to recognize hand poses and gestures. The Myo SDK provides default support for five gesture types: *double tap*, *fingers spread*, *wave right*, *wave left*, and *fist*. Also, Myo embeds a 9-axis Inertial Measurement Unit (IMU) that reports the orientation and movement of the wearing hand. Other technical details about the Myo armband, as well as previous work that employed it for various application scenarios, are presented in the side bar on this page.

Design and Implementation

Our prototype enables users to access TV content stored in their pockets and to display that content on the TV. According to the operation principle of the Smart-Pockets technique [18], the user points to a specific pocket (see Figure 2) and then points the hand towards the TV screen (Figure 1). The pocket is identified from the orientation of the hand delivered by the IMU unit embedded in the Myo armband and reported as quaternion data, $q = (x, y, z, w) \in R^4$. The action of pointing the hand in front of the body and towards the TV screen is detected using the same approach. To avoid selecting or changing the channel accidentally when the hand points in front of the body (*i.e.*, to avoid false positives), the user performs the default *double-tap* Myo gesture to confirm the command; see Myo's five pre-set gestures.[1] When the following series of three actions has been detected, the content associated to the pocket is displayed on the TV: (1) a pocket is identified using the orientation of the hand pointing to that pocket, (2) a pointing action is identified in front of the body, and (3) the *double-tap* gesture is performed. If the reverse sequence of actions is detected, the current TV channel is associated to the specific pocket pointed to by the hand. The side bar on the next page presents technical details about using the orientation of the hand to recognize the locations of pockets on the body using Myo's built-in IMU and a distance function that compares quaternions.

Configuration of Smart Pockets

Smart pockets can be easily configured by recording the orientation of the hand reaching to the locations of the pockets, such as the hand corresponding to reaching the shirt left pocket (see Figure 2, top) or the trousers rear right pocket (Figure 2, bottom). These two pockets can be easily accessed using the right hand. Other pockets can be easily set up, *e.g.*, an equivalent configuration of the shirt right pocket and trousers rear left pocket can be set up for the left hand as well.

POCKET DETECTION FROM ORIENTATION MEASUREMENTS OF THE MOVING HAND

The orientation data provided by Myo is reported in the form of unit quaternions $q_t = (w_t, x_t, y_t, z_t) \in [0, 1]^4$ at specific time intervals t. In order to use these data to identify absolute locations, quaternions need to be corrected by applying an offset with respect to a fixed, known location in space. This offset $q_{\text{offset}} = (w_{\text{offset}}, x_{\text{offset}}, y_{\text{offset}}, z_{\text{offset}})$, also a quaternion, is set once during the configuration step consisting in the user pointing their hand towards the TV screen. To apply the offset, each quaternion q_t is multiplied with q_{offset}, which actually corresponds to computing the rotation $q_r = (w_r, x_r, y_r, z_r)$ between quaternions q_t and q_{offset}, as follows:

$$w_r = w_{\text{offset}} \cdot w_t - x_{\text{offset}} \cdot x_t - y_{\text{offset}} \cdot y_t - z_{\text{offset}} \cdot z_t$$
$$x_r = w_{\text{offset}} \cdot x_t + x_{\text{offset}} \cdot w_t + y_{\text{offset}} \cdot z_t - z_{\text{offset}} \cdot y_t$$
$$y_r = w_{\text{offset}} \cdot y_t - x_{\text{offset}} \cdot z_t + y_{\text{offset}} \cdot w_t + z_{\text{offset}} \cdot x_t$$
$$z_r = w_{\text{offset}} \cdot z_t + x_{\text{offset}} \cdot y_t - y_{\text{offset}} \cdot x_t + z_{\text{offset}} \cdot w_t$$

We then compute the numerical distance between the offset-corrected orientation of the hand (q_{hand}) and the quaternion corresponding to that pocket location (q_{pocket}), a value recorded previously, using the following equation [7]:

$$d(q_{\text{hand}}, q_{\text{pocket}}) = 1 - < q_{\text{hand}}, q_{\text{pocket}} >^2$$

where $< q_{\text{hand}}, q_{\text{pocket}} >$ represents the inner product:

$$w_{\text{hand}} \cdot w_{\text{pocket}} + x_{\text{hand}} \cdot x_{\text{pocket}} + y_{\text{hand}} \cdot y_{\text{pocket}} + z_{\text{hand}} \cdot z_{\text{pocket}}$$

This result is then converted to an angle measurement:

$$\theta(q_{\text{hand}}, q_{\text{pocket}}) = acos\left(1 - 2\left(1 - < q_{\text{hand}}, q_{\text{pocket}} >^2\right)\right)$$

A Nearest-Neighbor classification algorithm is employed to compute the minimum distance across all the registered pockets and, thus, to identify the pocket that was reached by the hand.

DISCUSSION: TOWARDS EFFECTIVE GESTURE INPUT WITH LOW COGNITIVE EFFORT PERFORMED AT THE PERIPHERY OF ATTENTION

Our prototype contributes an important change of perspective for gesture-based control of TVs. While previous work strove to identify suitable gesture commands to effect various TV functions by means of gesture elicitation and agreement analysis [13, 16, 17, 19], the pointing gestures to smart pockets featured by our prototype are already intuitive, familiar, and easy to perform. From this perspective, there is an opportunity, to be explored and validated in future work, for implementing low-effort gesture-based interactions that take place at the periphery of user attention and, ideally, move outside the attentional field; see the continuum described by Bakker *et al.* [2, p. 6] for peripheral input. We believe that peripheral interaction [2] and, specifically, peripheral input based on gesture commands [6] fit well with the lean-back paradigm for television watching; see the side bar on the next page for a discussion about peripheral interaction. Consequently, Smart-Pockets pointing actions could be used to implement gesture input for the TV demanding a minimum level of attention, similar to that required by pressing buttons on the TV remote control. In turn, Smart-Pockets open up new research opportunities for low-effort gesture input performed with little to no cognitive load, in contrast to other types of gestures that, because of little agreement between users or low fit to the functions they execute [19], need to be learned for effective recall. Given the large interest in gesture-based input for the TV [1, 3, 4, 11, 13, 16, 17, 19], we hope that our explorations will contribute to and advance the practice of designing effective gesture user interfaces for television.

CONCLUSION AND FUTURE WORK

We presented a wearable prototype for controlling the content displayed on the TV by means of direct associations between preferred TV channels and physical pockets, *i.e.*, *TV channels are just another type of an object that we store in our pockets.* We positioned our implementation in the context of the recent interest and popularity received by wearable devices, to which end we presented a rapid prototyping approach using the Myo armband. Future work will consider user evaluations of the presented technique in the context of real-world applications, an investigation of the number of pockets and TV channels that users would like to control, as well as more technical work, for instance to make the Smart-Pockets interaction technique available on other types of wearables, such as smartwatches or smart rings [5]. Lastly, the opportunity to use Smart-Pockets to implement gesture-based peripheral input for the TV looks especially appealing.

ACKNOWLEDGMENTS

This work was supported by a grant of the Ministry of Research and Innovation, CCCDI-UEFISCDI, project no. PN-III-P3-3.1-PM-RO-CN-2018-0032 (3BM/2018), within PNCDI III.

PERIPHERAL INTERACTION

In this side bar, we briefly discuss peripheral interaction as well as implications for gesture-based control of the TV set. According to Bakker *et al.* [2], *"the interaction with computers will move to the periphery of attention. Inspired by the way we fluently divide our attentional resources over various activities in everyday life, we call this type of interaction peripheral interaction."* (p. 1). Peripheral interaction builds on top of ubiquitous computing, ambient intelligence, Internet of Things, context-aware computing, and considerate systems [2, pp. 3-5] to respond to the challenge of managing interactions with computing devices via fluent shifts between the user's center and periphery of attention.

A continuum for attentional resources

Bakker *et al.* [2] presented a continuum ranging from fully focused attention to tasks performed completely outside the attentional field, and identified three types of interactions: (1) in the *center of attention* (conscious, intentional precise control), (2) at the *periphery of attention* (intentional, direct imprecise control), and (3) *outside the attentional field* (subconscious and automatic).

Implications for gesture input for the TV

Our prototype makes use of a specific type of gesture: pointing to a pocket followed by pointing in the direction of the TV. We hypothesize that this familiar, intuitive, and easy to perform action can be executed at the periphery of attention and even outside the attentional field, which creates the premises for gesture-based control of the TV with little attention and, thus, little cognitive effort.

REFERENCES

[1] Gilles Bailly, Dong-Bach Vo, Eric Lecolinet, and Yves Guiard. 2011. Gesture-aware Remote Controls: Guidelines and Interaction Technique. In *Proc. of ICMI '11*. ACM, New York, NY, USA, 263–270. https://doi.org/10.1145/2070481.2070530

[2] Saskia Bakker, Doris Hausen, and Ted Selker. 2016. *Peripheral Interaction: Challenges and Opportunities for HCI in the Periphery of Attention* (1st ed.). Springer Publishing Company, Incorporated. https://doi.org/10.1007/978-3-319-29523-7

[3] Jan Bobeth, Johann Schrammel, Stephanie Deutsch, Michael Klein, Mario Drobics, Christina Hochleitner, and Manfred Tscheligi. 2014. Tablet, Gestures, Remote Control?: Influence of Age on Performance and User Experience with iTV Applications. In *Proc. of TVX '14*. ACM, New York, NY, USA, 139–146. https://doi.org/10.1145/2602299.2602315

[4] Nem Khan Dim, Chaklam Silpasuwanchai, Sayan Sarcar, and Xiangshi Ren. 2016. Designing Mid-Air TV Gestures for Blind People Using User- and Choice-Based Elicitation Approaches. In *Proc. of DIS '16*. 204–214.

[5] Bogdan-Florin Gheran, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2018. Gestures for Smart Rings: Empirical Results, Insights, and Design Implications. In *Proc. of DIS '18*. 623–635.

[6] Marigo Heijboer, Elise Hoven, Bert Bongers, and Saskia Bakker. 2016. Facilitating Peripheral Interaction: Design and Evaluation of Peripheral Interaction for a Gesture-based Lighting Control with Multimodal Feedback. *Personal Ubiquitous Comput.* 20, 1 (Feb. 2016), 1–22. https://doi.org/10.1007/s00779-015-0893-5

[7] Du Q. Huynh. 2009. Metrics for 3D Rotations: Comparison and Analysis. *Journal of Mathematical Imaging and Vision* 35, 2 (Oct 2009), 155–164. https://doi.org/10.1007/s10851-009-0161-2

[8] Heli Koskimäki, Pekka Siirtola, and Juha Röning. 2017. MyoGym: Introducing an Open Gym Data Set for Activity Recognition Collected Using Myo Armband. In *Proc. of UbiComp '17*. 537–546.

[9] Dong-Hoon Lee, Jung-Hyun Kim, Ho-Youn Kim, and Dong-Young Park. 2016. Remote Application Control Technology and Implementation of HTML5-based Smart TV Platform. In *Proc. of MoMM '16*. ACM, New York, NY, USA, 208–211. https://doi.org/10.1145/3007120.3007159

[10] North. 2018. Welcome to Myo Support. https://support.getmyo.com/hc/en-us

[11] Irina Popovici and Radu-Daniel Vatavu. 2018. Perceived Usability, Desirability, and Workload of Mid-Air Gesture Control for Smart TVs. In *Proc. of RoCHI '19*. 91–98. https://dblp.org/rec/bib/conf/rochi/PopoviciV18

[12] Jinfeng Rao, Ferhan Ture, Hua He, Oliver Jojic, and Jimmy Lin. 2017. Talking to Your TV: Context-Aware Voice Search with Hierarchical Recurrent Neural Networks. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17)*. ACM, New York, NY, USA, 557–566. https://doi.org/10.1145/3132847.3132893

[13] Kashmiri Stec and Lars Bo Larsen. 2018. Gestures for Controlling a Moveable TV. In *Proc. of TVX '18*. ACM, New York, NY, USA, 5–14. https://doi.org/10.1145/3210825.3210831

[14] Aaron Tabor, Scott Bateman, Erik Scheme, David R. Flatla, and Kathrin Gerling. 2017. Designing Game-Based Myoelectric Prosthesis Training. In *Proc. of CHI '17*. ACM, New York, NY, USA, 1352–1363. https://doi.org/10.1145/3025453.3025676

[15] Wan-Lun Tsai, You-Lun Hsu, Chi-Po Lin, Chen-Yu Zhu, Yu-Cheng Chen, and Min-Chun Hu. 2016. Immersive Virtual Reality with Multimodal Interaction and Streaming Technology. In *Proc. of ICMI '16*. 416–416.

[16] Radu-Daniel Vatavu. 2012. User-defined Gestures for Free-hand TV Control. In *Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12)*. ACM, New York, NY, USA, 45–48. https://doi.org/10.1145/2325616.2325626

[17] Radu-Daniel Vatavu. 2013. There's a World Outside Your TV: Exploring Interactions Beyond the Physical TV Screen. In *Proc. of EuroITV '13*. ACM, New York, NY, USA, 143–152. https://doi.org/10.1145/2465958.2465972

[18] Radu-Daniel Vatavu. 2017. Smart-Pockets: Body-deictic gestures for fast access to personal data during ambient interactions. *International Journal of Human-Computer Studies* 103 (2017), 1 – 21. https://doi.org/10.1016/j.ijhcs.2017.01.005

[19] Radu-Daniel Vatavu and Ionuţ-Alexandru Zaiţi. 2014. Leap Gestures for TV: Insights from an Elicitation Study. In *Proc. of TVX '14*. ACM, New York, NY, USA, 131–138. https://doi.org/10.1145/2602299.2602316