# Interaction Visualization for Analysing and Improving User Models

Stefan Lengauer
Lin Shao
Mariia Tytarenko
Graz University of Technology
Graz, Austria

Manfred Klaffenböck
TU Wien
Vienna, Austria

Tobias Schreck
Graz University of Technology
Graz, Austria

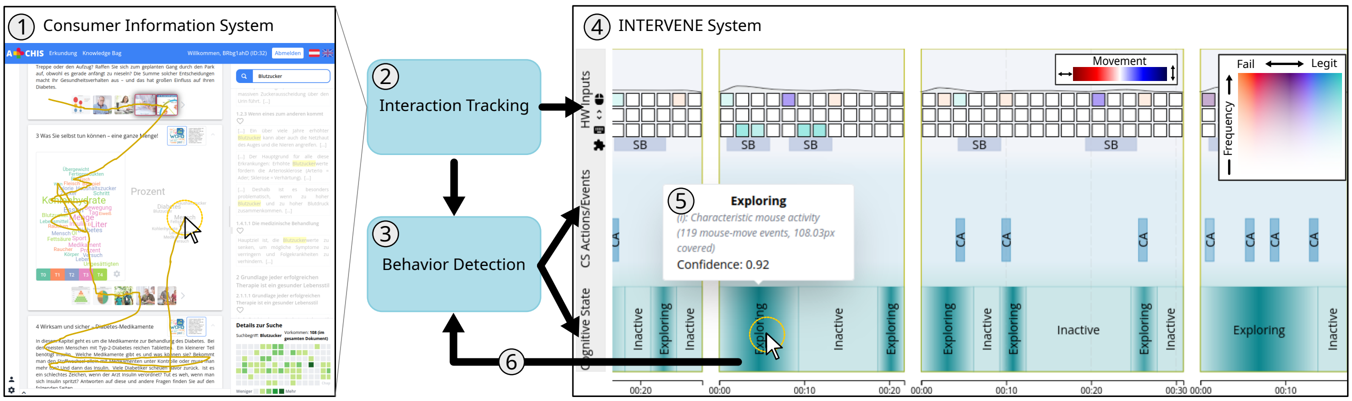


**Figure 1: Our INTERactiVe bEhavior aNalysEr (INTERVENE) ④ illustrates a user's interaction with a web-based information system ①. To this end, all interactions (mouse and keyboard events) are precisely tracked ② and fed into models for behavior detection ③. The INTERVENE system allows experts to analyse and validate these models by relating them to the co-occurring interactions and on-demand insights ⑤. Finally, if miss-classifications are recognized, these can be annotated and fed back to improve the models ⑥.**

## ABSTRACT

Many web-based systems such as online retail, information systems or search engines track the interactions users have with them. Tracked data can comprise high-level information like dwelling time, reviewed items, and clicked elements, but also fine-grained information in the form of mouse trajectories and keystrokes. While these data are often fed into user- or behavior models in recommender systems, there are few approaches for interactive visual exploration of multi-modal and complex interaction patterns. Yet, the thorough analysis could reveal important insights for the design and evaluation of said models. We propose a suitable visual analysis approach that allows to validate and correct models in an intuitive and interactive manner. Our tool provides insights into concrete user (inter)actions and also estimates more complex behavioral patterns. Level of detail views in our system outlines the certainty of detected behaviors and serve the explainability. Our approach can help engineers to understand user interactions and improve behavioral models.


*UMAP Adjunct '24, July 01–04, 2024, Cagliari, Italy*

ACM ISBN 979-8-4007-0466-6/24/07
https://doi.org/10.1145/3631700.3664877

## CCS CONCEPTS

• **Computing methodologies** → **Model development and analysis**; **Model verification and validation**; • **Information systems** → *Information systems applications*.

## KEYWORDS

User Modelling, Model Validation and Analysis, Interaction Visualization

**ACM Reference Format:**
Stefan Lengauer, Lin Shao, Mariia Tytarenko, Manfred Klaffenböck, and Tobias Schreck. 2024. Interaction Visualization for Analysing and Improving User Models. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP Adjunct '24), July 01–04, 2024, Cagliari, Italy*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3631700.3664877

## 1 INTRODUCTION

Many modern web-based information systems collect an abundance of data pertaining to users' interaction with the system. Data gathered this way can be used to establish user models, which in turn can be utilized for developing recommendation strategies [20]. Such user tracking has also arrived in health information applications [19]. Users of such systems might have special assistive needs [7] or at least benefit from accessibility features (e.g., elderly people) they are not aware of [23]. Inputs provided by those users might not always be meaningful (e.g., clicking into non-clickable areas) and therefore not lead to adequate results. Considerations like these increase the complexity of these possibly already complex, often ML-based, user models. Therefore, both the validation and the improvement of such models can become increasingly difficult. However, it has been shown [11, 18] that human-in-the-loop approaches with interactive interfaces are able to support the understanding of model uncertainty, improve prediction accuracy and even teach concepts that are otherwise hard to learn.

We propose a multimodal feedback interface – INTERactiVe bEhavior aNalysEr (INTERVENE) – for visualizing users' interactions with an information system which, to the best of our knowledge, has not been used in the context of user interaction analysis. Relying on a stack of different user processes [17], we employ fundamental empirical models to showcase the capabilities of our visualization to provide insight into how our models arrive certain predictions and which input data was responsible. That is, we implement a precise tracking to an existing information system (Sec. 3.1), which captures hardware inputs (mouse and keyboard events) together with the context they occur in. Context describes the elements with whom users interacted and potential system responses. E.g., mouse clicks carry entirely different meaning depending on whether they were performed on a clickable button or on unresponsive background elements. The INTERVENE tool shows all this comprehensively in a tiered timeline visualization with custom-built abstraction and numerous detail-on-demand components (Sec. 3.3). Our main contribution is a novel multi-modal interface for the visual analysis of user interactions, enabling overview and comparison between low-level inputs and high-level abstractions of user behavior.

## 2 BACKGROUND AND RELATED WORK

Visual Analytics combines model-based analysis with interactive visualization to accelerate insight generation [13], but it can also facilitate the design and refinement of user models [3, 24]. Our visualization approach builds upon an existing behavior model [17] to support the development of models for adaptive systems and recommendation strategies.

Leveraging visualization for model development has a long history [10]. Various interfaces for validating and enhancing user models have been proposed [11, 18]. Our approach specifically focuses on visualizing the user interaction history, which can be helpful in gaining a better understanding of the user as well as designing and (re-)evaluating personalized adaptation strategies [24]. Several approaches have been proposed for the visualization of user interaction histories. Heer et al. [15] used a comic-strip metaphor to show the different visualizations a user created. This works well where each interaction results in an interesting visual state but is less suited for more general user interactions. This approach also does not capture the time passed between the represented states. Carrasco et al. [9] used animated bubble charts to visualize the browsing history of users. This approach is suited for visualizing a single user's history for personal reflection, but it does not perform well when it comes to overview and comparison of multiple entities. Guo et al. [13] visualized multiple aspects of interest (like select, explore, and elaborate) on a single timeline. This does not work well when those aspects can happen simultaneously. For instance, exploration and elaboration would not necessarily have to be mutually exclusive. We follow an approach that has its origin in film and multimedia analysis [5, 8], where multiple stacked timelines, each representing a different aspect of interest, are utilized. This approach provides a better overview and allows for an easy comparison between the different aspects. Our method employs such tiered timelines in order to connect observable low-level interactions (e.g., keystrokes or mouse inputs) with high-level assumptions about a user's cognitive state.

Recommender systems based on user interaction history have been used in various application domains in order to make recommendations related to movies [1], visualizations [2], locations [22], and sports [12], just to name a few. Our current focus is on health information systems, but it could be easily extended to other domains as well.

## 3 THE INTERVENE SYSTEM

The input for our INTERVENE system is interaction data collected from users exploring a web-based information system, which we collect with a dedicated tracking system (Fig. 1②). Those are fed into models for detecting behavioral patterns and cognitive states (Fig. 1③). Our proposed design visualizes both interaction data and recognized states in a coherent manner (Fig. 1④).

### 3.1 Interaction Tracking

Different solutions exist for the tracking of mouse, touch, keyboard, and other low-level hardware inputs in web environments. Commercial products[1] combine this tracking with different analysis tools, such as heatmapping or journey analysis. But there are also research-oriented tools [4], which insert JavaScript code for tracking via a PHP proxy. However, for our experimental system, we rely on a self-made tracking due to the complexity of the used information system (various expandable/scrollable elements) and the fact that we also want to save the context of an interaction (e.g., which element was in-focus at the moment of the action or which element was hovered by the mouse). Both requirements pose a substantial challenge to existing tracking tools.

In our setup, we store the $x, y$-position of the mouse within the view, together with context in the form of the visual element at this location and events triggered by that. The sampling frequency of this information depends on the browser implementation, but in our development setup, we observed frequencies of up to 25 Hz. We detect and save all key-presses and releases and the element which was active at the respective times. Also, dwelling times of all visual components of the information system, together with the respective appearance at the time, are captured.

### 3.2 Behavior Modelling

Regarding behavioral models, we follow the hierarchic processes/events landscape outlined in Lengauer et al. [17], who describe increasingly complex interaction patterns based on the raw hardware inputs. They distinguish between two tiers of processes. Firstly, *Context-Sensitive Action/Events (AEs)*, which are actions defined by hardware inputs but requiring a specific execution context. E.g., the display of a thumbnail if the mouse moves over a hoverable element. Secondly,

[1] https://contentsquare.com/clicktale/

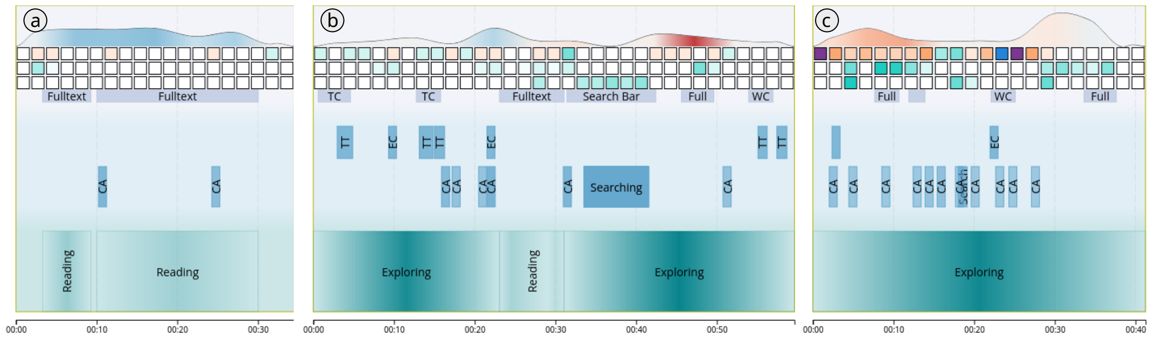


**Figure 2: Selected interaction examples of about one minute, characteristic for ⓐ *Reading*, ⓑ *Exploring*, and ⓒ *Confused* CSs.**

they define the *Cognitive States (CSs)*: (i) *Pause/Inactive*, (ii) *Reading/Inspecting*, (iii) *Examining/Exploring*, and (iv) *Lost/Confused*.

From this taxonomy, our current prototype features the AEs: *Click-fail*, *Typing-fail*, *Typing*, *Expand/Collapse*, *Hover*, *Search*, and *Conscious Action*, as well as the CSs: *Inactive*, *Reading*, and *Exploring*. We employ a rule-based strategy to detect those states. For *Click-fails*, we verify that the clicked elements belong to a pre-determined list of clickable elements. For *Type-fails*, we check if the currently focused element permits keyboard input. *Expand/Collapse* and *Hover* are tracked via callbacks on the respective elements. *Search* is defined as a series of keyboard inputs while a search field is in focus. A *Conscious Action* is characterized by a mouse trajectory to a clickable element where a click is performed [16]. The underlying probability function for such an event is based on the "straightness" of the mouse trajectory and the movement speed in the moments preceding the click.

For the Cognitive States (CSs) we developed a cascading rule set. (1) If the mouse dwells on an element displaying floating text, we assume a *Reading* state. The probability of this estimate increases if the user follows the lines with the cursor in characteristic reading patterns [16]. To this end, we implement a straightforward criterion based on the swept line width and the movement in a vertical direction. (2) If no hardware input happens for some time, an *Inactive* state is assumed. (3) For any other case, we assume an *Examining* state and estimate its likelihood, based on a simple activity model. The fourth CS – *Lost/Confused* – is challenging to distinguish from "Examining" using a rule-based approach and is therefore not implemented yet. Please note that the presented rules serve as an empirically developed baseline designed to demonstrate our visualization approach. These relatively simple rules can easily be replaced with more sophisticated models.

### 3.3 Visual Design

The visual design of the INTERVENE system (Fig. 1,④) is a time axis along which a user's exploration sessions are displayed as blocks. The vertical dimension projects the different complexity tiers of user interactions in respectively colorized segments. The uppermost tier for the hardware inputs shows the frequencies and "success"-rates of mouse and keyboard usage. I.e., for the mouse input, the overall mouse movement (covered pixel per second) as well as its current directivity (vertical vs. horizontal) is visualized through an appropriately colorized area chart. A bluish hue indicates an overwhelmingly horizontal mouse movement and a reddish vertical movement. Below that are three rows, illustrating mouse clicks, -scrolls and keyboard inputs with a tilebar [14] metaphor. From those, we can perceive both periods of high usage (saturation) and effectiveness (hue), since different inputs are oftentimes not supported by an element. The line below the tiles reflects the dwelling time on the information system's different visual components. In our example, things like *Fulltext*, *Search Bar*, *Word Cloud (WC)*, and *Topic Cloud (TC)*. The second tier comprises all Context-Sensitive Action/Events (AEs) with a vertical ordering according to their complexity. If the occurrence of an AE is inherently ambiguous, e.g., for a *Conscious Action*, the respective occurrence confidence (as determined by our model) is encoded in the blocks' opacities. The third and final tier of Cognitive States (CSs) makes up the bottom row with the opacity of a block similarly reflecting the confidence of a state estimate as returned by the models.

## 4 EXPERIMENTAL USE CASE FROM THE HEALTH DOMAIN

To demonstrate the usability of our visual design, we present initial experimental results for an Adaptive Consumer Health Information System ($A^+$CHIS) on type 2 diabetes [21] (Fig. 1①). This system is designed for non-medical users and aims to provide adaptive, interactive, and visual information based on users' exploration and navigation patterns. To this end, the $A^+$CHIS employs different visual components, implementing text abstraction methods like word clouds, topic models, and the like, to visualize medical information extracted from respective documents. Dedicated components also display pictorial content, like infographics and provenance tools allow users to see their exploration history. For our experiment, the brochure *Den Diabetes im Griff* [6] by the German health insurance provider AOK is used. This document spans over 130 pages and provides extensive health information, including figures, tables, and infographics. An exemplary exploration process with the $A^+$CHIS system and subsequent interaction analysis with our INTERVENE is showcased at https://youtu.be/DIHGH0CINMY. Selected example interactions are also shown in Fig. 2, with ⓐ showing a *Reading* state with constant horizontal sweeping mouse movement over a text display; ⓑ exhibiting typical exploration behavior characterized by high overall activity and the presence of many CAs; ⓒ showing roughly similar patterns but a closer inspection using the detail-on-demand functionalities (Fig. 1④) reveals many erroneous clicks and overall irregular mouse movements. Even though this is labeled as *Exploration* behavior by our naïve model, this is more indicative of a *Confused* state and constitutes a prime example where an expert user could relabel this estimate to improve upon the model.

## 5 DISCUSSION

As our system is a work in progress, there are several ideas for further development. That is, at the current state, the INTERVENE system is tightly coupled to the above mentioned A$^+$CHIS for which we implemented a tailored tracking mechanism. More specifically, the system has a range of canvas and SVG elements to display interactive visualizations that we track. We tried, however, to generalize the observed interactions as well as possible. I.e., interactions like expanding/collapsing components, hovering, or searching are possible on most web-based information systems. Hence, an obvious future work item is to decouple the tracking from this example system and generalize it such that it can be readily used for any system. E.g., by injecting it via a proxy [4]. Also, at the current state, the INTERVENE is only fit for the analysis and validation of behavioral models. I.e., the option to provide feedback regarding erroneous classifications is not yet implemented. In the next extension, it is planned that a user can either confirm detected behaviors or reclassify them, informed by observing the mouse movements and other interactions at the respective time. This information will then be fed back to the behavior model, which we will change from a rule-based model to an online learning concept.

## 6 CONCLUSION

INTERVENE is a novel tool which displays users' interaction with a system in an intuitive and comprehensible manner. It aids an expert to analyze, validate, and – eventually – improve behavioral models. We demonstrate the capabilities of our concept through a working prototype implementation and a selected use case scenario from the health domain.

## ACKNOWLEDGMENTS

This work was funded by the FWF as part of the project 'Human-Centered Interactive Adaptive Visual Approaches in High-Quality Health Information' (A$^+$CHIS; Grant No. FG 11-B).

## REFERENCES

[1] Charu C Aggarwal et al. 2016. *Recommender systems*. Vol. 1. Springer, Yorktown Heights, NY, USA. https://doi.org/10.1007/978-3-319-29659-3

[2] Panayiotis Andreou, Christos Amyrotos, Panagiotis Germanakos, and Irene Polycarpou. 2023. Human-centered Information Visualization Adaptation Engine. In *Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization* (Limassol, Cyprus) *(UMAP '23)*. Association for Computing Machinery, New York, NY, USA, 25–33. https://doi.org/10.1145/3565472.3592961

[3] N. Andrienko, T. Lammarsch, G. Andrienko, G. Fuchs, D. Keim, S. Miksch, and A. Rind. 2018. Viewing Visual Analytics as Model Building. *Computer Graphics Forum* 37, 6 (2018), 275–299. https://doi.org/10.1111/cgf.13324

[4] Ernesto Arroyo, Ted Selker, and Willy Wei. 2006. Usability tool for analysis of web designs using mouse tracks. In *CHI'06 extended abstracts on Human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 484–489. https://doi.org/10.1145/1125451.1125557

[5] Olivier Aubert and Yannick Prié. 2007. Annotations de documents audiovisuels-Temporalisation et spatialisation. In *Atelier Interfaces pour l'annotation et la manipulation d'objets temporels: une comparaison des outils et des paradigmes dans le domaine musical et cinématographique*. Paris, France, 1–3.

[6] J. Baumgart, U. Viegener, and C. Pohl. 2021. Den Diabetes im Griff: Ein Handbuch für Patientinnen und Patienten mit Diabetes mellitus Typ 2. *AOK-Bundesverband, Berlin* (2021). https://www.aok.de/pk/magazin/cms/fileadmin/pk/pdf/patientenhandbuch-diabetes.pdf

[7] Emily C Bouck. 2015. *Assistive technology*. Sage Publications, Los Angeles, USA.

[8] Hennie Brugman, Albert Russel, and Xd Nijmegen. 2004. Annotating Multi-media/Multi-modal Resources with ELAN.. In *LREC*. Lisbon, 2065–2068.

[9] Matthew Carrasco, Eunyee Koh, and Sana Malik. 2017. popHistory: Animated Visualization of Personal Web Browsing History. In *Proceedings of the 2017 CHI Conference (CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 2429–2436. https://doi.org/10.1145/3027063.3053259

[10] Andrew W. Crapo, Laurie B. Waisel, William A. Wallace, and Thomas R. Willemain. 2000. Visualization and the process of modeling: a cognitive-theoretic view. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Boston, Massachusetts, USA) *(KDD '00)*. Association for Computing Machinery, New York, NY, USA, 218–226. https://doi.org/10.1145/347090.347129

[11] John J. Dudley and Per Ola Kristensson. 2018. A Review of User Interface Design for Interactive Machine Learning. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 8 (jun 2018), 37 pages. https://doi.org/10.1145/3185517

[12] Ciara Feely, Brian Caulfield, Aonghus Lawlor, and Barry Smyth. 2023. Modelling the Training Practices of Recreational Marathon Runners to Make Personalised Training Recommendations. In *Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization* (Limassol, Cyprus) *(UMAP '23)*. Association for Computing Machinery, New York, NY, USA, 183–193. https://doi.org/10.1145/3565472.3592952

[13] Hua Guo, Steven R. Gomez, Caroline Ziemkiewicz, and David H. Laidlaw. 2016. A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights. *IEEE TVCG* 22, 1 (2016), 51–60. https://doi.org/10.1109/TVCG.2015.2467613

[14] Marti A Hearst. 1995. Tilebars: Visualization of term distribution information in full text information access. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., USA, 59–66. https://doi.org/10.1145/223904.223912

[15] Jeffrey Heer, Jock Mackinlay, Chris Stolte, and Maneesh Agrawala. 2008. Graphical Histories for Visualization: Supporting Analysis, Communication, and Evaluation. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1189–1196. https://doi.org/10.1109/TVCG.2008.137

[16] Jeff Huang, Ryen White, and Georg Buscher. 2012. User see, user point: gaze and cursor alignment in web search. In *Proc. ACM Human Factors in Computing Systems (CHI)*. 1341–1350.

[17] Stefan Lengauer, Michael Bedek, Cordula Kupfer, Lin Shao, Dietrich Albert, and Tobias Schreck. 2023. Recognizing User Behavior from Interactions for Adaptive Consumer Information Systems. In *International Conference on Interactive Media, Smart Systems and Emerging Technologies (IMET)*. The Eurographics Association. https://doi.org/10.2312/imet.20231251

[18] Swati Mishra and Jeffrey M Rzeszotarski. 2023. Human Expectations and Perceptions of Learning in Machine Teaching. In *Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization* (Limassol, Cyprus) *(UMAP '23)*. Association for Computing Machinery, New York, NY, USA, 13–24. https://doi.org/10.1145/3565472.3595612

[19] Maurice Mulvenna, Raymond Bond, Alexander Grigorash, Siobhan O'Neill, and Assumpta Ryan. 2018. HILDA - A Health Interaction Log Data Analysis Workflow to Aid Understanding of Usage Patterns and Behaviours. In *Unknown Host Publication*. Society for the Study of Artificial Intelligence and Simulation of Behaviour, United Kingdom.

[20] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2021. Recommender systems: Techniques, applications, and challenges. *Recommender systems handbook* (2021).

[21] Lin Shao, Stefan Lengauer, Hossein Miri, Michael A. Bedek, Bettina Kubicek, et al. 2023. Visual Document Exploration with Adaptive Level of Detail. In *18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications: VISIGRAPP 2023*, Vol. 3. SciTePress-Science and Technology Publications, 133–141. https://doi.org/10.5220/0011621800003417

[22] Qinyong Wang, Hongzhi Yin, Tong Chen, Zi Huang, Hao Wang, Yanchang Zhao, and Nguyen Quoc Viet Hung. 2020. Next Point-of-Interest Recommendation on Resource-Constrained Mobile Devices. In *Proceedings of The Web Conference 2020* (Taipei, Taiwan) *(WWW '20)*. Association for Computing Machinery, New York, NY, USA, 906–916. https://doi.org/10.1145/3366423.3380170

[23] Jason Wu, Gabriel Reyes, Sam C. White, Xiaoyi Zhang, and Jeffrey P. Bigham. 2021. When can accessibility help? an exploration of accessibility feature recommendation on mobile devices. In *Proceedings of the 18th International Web for All Conference* (Ljubljana, Slovenia) *(W4A '21)*. Association for Computing Machinery, New York, NY, USA, Article 21, 12 pages. https://doi.org/10.1145/3430263.3452434

[24] Kai Xu, Alvitta Ottley, Conny Walchshofer, Marc Streit, Remco Chang, and John Wenskovitch. 2020. Survey on the Analysis of User Interactions and Visualization Provenance. *Computer Graphics Forum* 39, 3 (2020), 757–783. https://doi.org/10.1111/cgf.14035