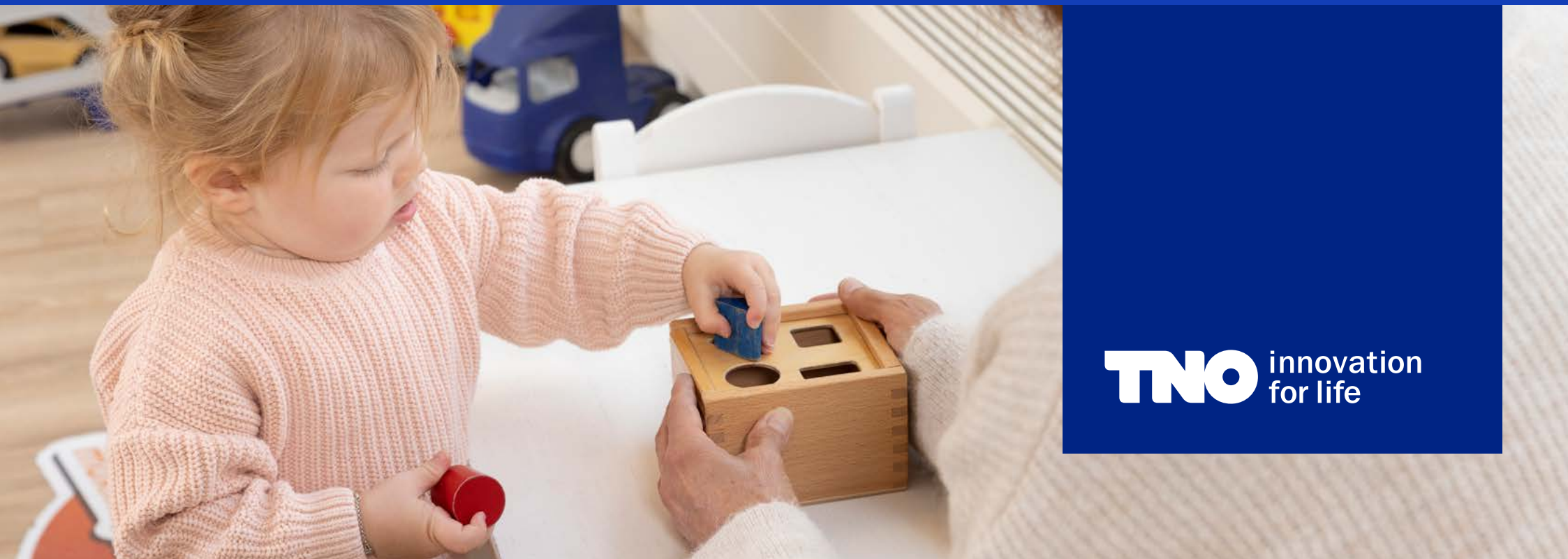
Impact story

Tracking children's development worldwide with the D-score



TNO innovation for life

What is an impact story?

It can take anywhere from 5 to 20 years for a technological innovation to actually make a difference. At TNO, we know this all too well. Equally, it takes a lot of capital and expertise to get from idea to application. While not every innovation journey leads to a successful market launch, there are some highly inspiring success stories. These include technological innovations that are not only profitable, but more importantly make a positive impact on society. But that takes a lot of patience and perseverance. How does this work in practice? In this series of Impact Stories, TNO employees and partners look back on how they developed technological innovations that have clearly proven themselves in the long run.

How do you visualise children's development clearly, reliably, and consistently? And how can you do this globally to compare children's development across different countries? There are major parties working on this around the world, but so far it is TNO that seems to have come up with the best approach – with a unit of measurement called the D-score. In 2018, the World Health Organisation (WHO) called this development promising, stating it had a high chance of being adopted globally as a vital tool for monitoring child development. It all started some 20 years ago when Stef van Buuren, Principal Scientist in the Child Health department at TNO, published a statistical principle on existing child development data purely out of curiosity.

200 million children not reaching their developmental potential

A child's first 1,000 days are crucial for their cognitive, emotional, and physical development. Children who don't receive adequate nutrition, protection, and stimulation at this early stage of life can suffer the consequences throughout their lives. In 2011, medical journal The Lancet published a series of articles on early childhood development (ECD). One of these articles reported that it was estimated that around 200 million children in the world – mainly in low and middle-income countries – were not reaching their developmental potential. Poverty, malnutrition, lack of education, and poor health all negatively affect children's brain development and, in turn, their future opportunities. This is because developmental delays often make it impossible to keep up in school and earn a good living later on in life. So, not only are their future prospects at stake, but if so many children are left behind, this also has a negative impact on a country's economic development and level of well-being.

Urgent call

The article in The Lancet also contained a call to action. A global coalition was needed to champion policies, investments, and interventions that would help improve care, nutrition, and education for younger children. This plea for immediate global action also called for investment in the first years of children's lives to give them a fair start and reduce inequality, especially in poorer regions.

Wanted: a new measurement methodology

While diligently argued, the article in The Lancet caused a stir. How had these 200 million children fallen under the radar of major global organisations such as WHO and Unicef? After much analysis and discussion, it soon became clear: there had to be a new measurement methodology. One that could depict global child development consistently. Not only would this quickly identify individual disadvantages in that field, but it would also help compare local areas, cities, regions, countries, and continents.

The author of that article was none other than world-renowned 'missing data professor',
Stef van Buuren. Besides his work for TNO, Stef is also Professor of Statistical Analysis of Missing Data at Utrecht University (or simply 'Professor of Missing Data', as he is informally known both behind the scenes and in the outside world). He now enjoys global fame as the creator and disseminator of the MICE algorithm: a powerful tool for analysing incomplete data that thousands of researchers worldwide use every year. So, he was quite used to delving into complicated statistical problems. This eventually led to his discovery of a new methodology now known as the D-score, which has also garnered a great deal of international attention.

Not the first plea for a universal measure

Let's go back to 2006, when a scientific article was published arguing for a universal measure of child development. This should, the author wrote, be broadly similar to the universal scale for length measurements. He also immediately presented a concept for a statistical model that could capture a child's development in a single number.

A clever idea

Now a little further back in time to 1989 and 1990, when TNO conducted a large study on child development in the Netherlands. That research focused on the physical and psychomotor development of children in the first years of their life. The study included some 2,000 children and its results were central to improving youth healthcare. "On top of that, the researchers did something incredibly clever at the time," says Van Buuren. "They not only looked at things that were normal for children to do at their age. But they also looked at the extent to which those children could do things they were not normally expected to do until later in their developmental growth. This allowed them to visualise not only which children were lagging behind, but also which children were actually ahead in their development."

Eureka moment

In terms of child development, TNO had a very reliable dataset. Stef also knew this. Fast-forward again to 2006 and he suddenly had a brainwave. What would happen if he ran that data through the Rasch model? This is an existing statistical method originally developed to quantify individual performance in tests or questionnaires. Stef took the test straight away. “I started with a graph where all the data points were criss-crossed. After applying the statistical model, I saw a graph with equivalent curved lines all running neatly side by side. I had never seen a solution with such a beautiful outcome. It looked fantastic. Plus, the Rasch model is an incredibly strict one. You don’t just get it to fit that easily. That was a real eureka moment.”

Nine ways the D-score can be used:

1. **Early intervention programmes**: the D-score reveals which children are developmentally delayed at a young age, making it possible to make appropriate interventions and deploy support programmes at an earlier stage, such as pre-school programmes or special education.
2. **Longitudinal development studies**: by measuring the D-score at multiple points in time, you can measure children’s development over longer periods of time, for example to identify growth patterns or changes within the population.
3. **National and international comparisons**: if several countries were to map out child development at national level, specific aspects can then be compared at global level. This, in turn, helps to understand the influence of various environmental factors, such as poverty or access to healthcare. But the D-score is equally useful for studying the regional differences within a single country.
4. **Predicting school performance**: in early childhood studies, the D-score can help predict how a child will do in school later on.
5. **Assessing interventions**: the D-score can be used to evaluate the effectiveness of development programmes or pedagogical interventions. This can be done by determining children’s D-scores before and after an intervention.
6. **Guidance for educators**: parents and caregivers can be offered personalised guidance based on their child’s D-score, for example advice or activities designed to stimulate specific areas of development.
7. **Diagnosing developmental disorders faster**: by combining the D-score with other diagnostic tests to identify children with developmental disorders, such as autism or ADHD, a diagnosis can be given more quickly and treatments can be started earlier.
8. **Comparing measurements taken with different instruments**: previously incomparable measurements can be converted to the same scale when determining the D-score, so comparisons can be drawn between children, interventions, and countries.
9. **Adaptive measurement**: the D-score also makes it possible to tailor a measurement at an individual child level.

“I started with a graph where all the data points were criss-crossed. After applying the statistical model, I saw a graph with equivalent curved lines all running neatly side by side (...) That was a real eureka moment.”
Stef van Buuren - Principal Scientist at TNO

Underwhelming response

Later that year, Stef shared his findings in a scientific paper, outlining the statistical model needed to produce a universal measure of child development. He was anticipating a great deal of interest, not just from global child welfare organisations, but also from policymakers and youth care providers. After all, the model would allow them to build a quick picture of how children’s development is progressing and whether there might be any delays. Stef braced himself for the doubtless high level of interest he would receive from the field. But none of that all. “Together with TNO colleagues, I investigated the principle further and looked at how much it holds up in practice. We then wrote another scientific paper on it, which was published in a paediatric journal in 2007, but that too wasn’t picked up at all. Well that’s a shame, I thought at the time. And that was that. I thought maybe a prince will come along one day and awaken this sleeping beauty.”

Top statistician wanted

That ‘prince’ was still another seven years away. In 2014, Stef received an email from Tim Cole, an innovative statistics professor in England who he knew well and he had previously worked with on child growth models. Tim told him about an upcoming meeting of the Global Child Development Group (GCDG), where leading experts from around the world were seeking a top statistician to help develop a global measurement tool for child development. Tim knew just the man for the job and had already passed on his name.

Trip to Jamaica

Stef headed off to the GCDG’s base in Kingston, Jamaica for a three-day meeting with 13 child development experts. The meeting was held in a hotel basement: no windows, no daylight, no distractions. The expert group was led by Maureen Black, a renowned US psychologist. She had also previously collaborated on studies exploring solutions for the 200 million children unable to reach their developmental potential.

Exactly what we’re looking for!

The 13 experts all knew each other well. Stef was not just the only statistician there, but also the only man. At the start of the meeting, he felt a bit of an outsider. But that feeling soon disappeared once he started explaining how he had arrived at a new statistical methodology based on TNO data on Dutch children’s development. There was no reason to suggest the same principle couldn’t be applied worldwide, he added. With that, he immediately had the full attention of the experts. “This is exactly what we are looking for!” Black concluded at the end of the first day.

Mapping out different measuring instruments

The experts' laptops contained a wealth of data: not only on child development in different countries, but also on what measurement tools are used around the world. Over the next two days of the meeting, they carried out an inventory together with Stef. By the end of the last day, they had a spreadsheet that mapped out the overlap between 16 different instruments used to measure child development. That gave Stef and his TNO colleagues a firm starting point for the global D-score he had so clearly envisioned several years earlier.

A recipe for calculating child development

Can a child walk? Can they say three words yet? Do they smile back at you? Can they stack two blocks? These are some examples of child development milestones mapped in the Netherlands under the Van Wiechen scheme. Based on the physical, emotional, and communication milestones recorded, the D-score can then summarise a child's development in a single number. This works not only in the Netherlands: once calibrated, the D-score can also be used with child development data collected in other countries through other measurement instruments. So, the D-score does not replace measurement instruments, but is a rather like a recipe for calculating easy-to-understand developmental scores. This makes it possible to compare measurements taken by different instruments, both individually and for groups. The D-score also allows comparisons between cities, regions, countries, and continents. While it is not quite being used at that scale yet, it is certainly heading in that direction.

International collaboration

Finally, Stef had got the ball rolling. Now all he needed was some funding, but Stef knew what to do. Shortly before his trip to Jamaica, he was asked by the Bill & Melinda Gates Foundation (BMGF) to help think about the best way to portray a child's first 1,000 days. Together with the GCDG group, he submitted proposals to the BMGF and Bernard van Leer Foundation. and they loved them. Of course, the research subsequently made possible thanks to the funding was not done alone, but with the help of TNO colleagues and a group of international researchers and specialists, including the experts he had got to know so well at the Jamaica meeting. During that research, a huge amount of data from 11 countries was gathered and harmonised. Next, the Rasch model had to be extended. All to produce the first version of an overall D-score.

Three initiatives, one goal

Meanwhile, there were two other initiatives underway: WHO experts and a group of specialists from Harvard were also working on child development measurement tools at the time, each also hoping to come up with a universal outcome measure. Awkward. After all, they were all working towards the same goal and their work overlapped in many areas, so it didn't make much sense to have three similar initiatives running side by side. Not least because developing a universally applicable measurement methodology takes an enormous amount of time and money. So, WHO decided to bring the three groups together to really start working together towards the common goal.

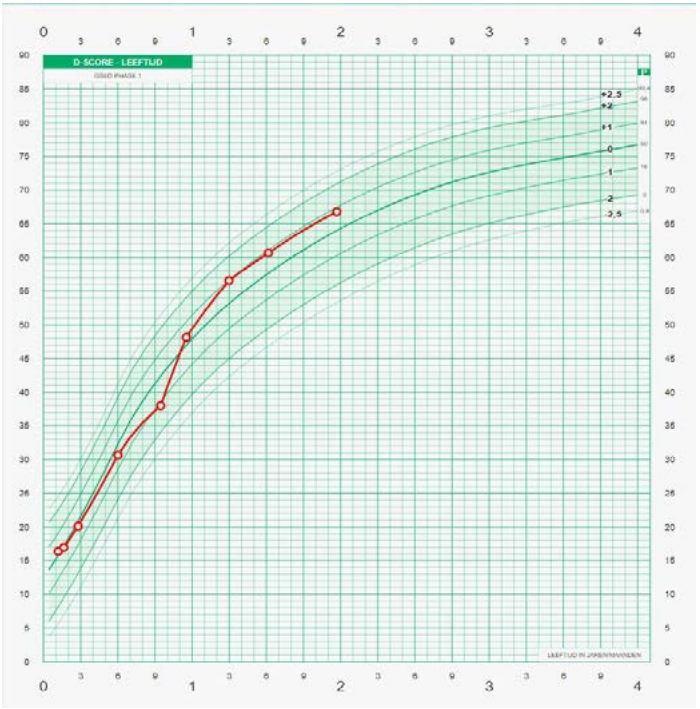
Crunch time

In 2018, the time had come to make a decision. Which parts of the three

"My idea was to show how the D-score worked through web applications (...) That would give more insight into the method, where it would also immediately become clear that a lot of multifaceted preliminary research had preceded it and that a solid foundation was already in place."
Iris Eekhout, Statistician at TNO

Winner of the H.A. Lorentz Prize

On 14 December 2023, Stef van Buuren received the H.A. Lorentz Prize for his groundbreaking applied scientific research that has had an extraordinary impact on society. He received the prestigious TNO award for his years of pioneering work on universal standards to measure and monitor children's growth and development worldwide. A few months earlier, a letter of recommendation had arrived at TNO from the US, sent by a certain Professor Maureen M. Black. The child psychologist whom Van Buuren first met at the meeting in Jamaica was instrumental in ensuring that the D-score gained the global attention that it since has.



The D-score summarises a child's development in a single number.

WHO, Harvard, and TNO initiatives were most relevant? On 1 November, WHO convened a meeting in Washington DC of 25 international experts from the likes of Harvard University, John Hopkins University, Stanford University, and TNO to decide. Those experts included delegates from WHO, financial backers, statisticians, and subject matter experts on children, such as paediatricians, psychologists, and health researchers.

Clear ideas

No-one likes a slideshow presentation that's just about the formulas and numbers. The D-score team at TNO was all too aware of this, and so was Iris Eekhout. A year after completing her PhD on missing data, a subject that Stef van Buuren is a global authority on, she noticed a vacancy in Stef's department at TNO. In late 2015, Iris joined as a statistician. About a year later, as Stef's D-score was gaining momentum, it became the topic she started spending most of her time working on at TNO. She already knew the D-score inside out when the chance arose to present it in Washington in 2018. Straight away, she had clear ideas on how the TNO team could best tackle it.

Don't say it, show it!

"We had to give the audience a kind of live demonstration, so they could immediately grasp how thorough the method was, without overwhelming them with complex information and formulas," Iris recalls. "My idea was to show how the D-score worked through web applications. And these were applications they could use themselves. That would give more insight into the method, where it would also immediately become clear that a lot of multifaceted preliminary research had come before it and that a solid foundation was already in place. So, we went with a visual presentation, where we could clearly show right there and then that the model cleverly combines different data from different measurement instruments to produce one easily interpretable score. Therein lies the true power of the D-score."

Interestingly, the web applications Iris mentioned didn't even exist at the time. So, she built them herself. It was a huge amount of work, but she knew it would be worth it.

The D-score presentation

Then it was game on. Stef and Iris took turns on stage in Washington to explain the theory and provide insights into the model and results through web applications. The presentation went without a hitch. They engaged with their audience. The web applications worked a

treat. And Stef was able to tackle all the critical questions from the few statisticians present in the room in a way that was clear to all.

Wonderful moment

The experts present were clearly impressed by how the D-score could be applied in practice. The majority of experts even said that the D-score might end up becoming a global standard. “That was a wonderful moment,” Iris looks back. “When we got back to the Netherlands, we were still in a kind of daze.”

A big boost, thanks to WHO

The enthusiasm for the D-score provided was wonderful news. And that was the point when it really took off. Because the main goal of the international collaboration was to introduce an extensively tried-and-tested measurement tool that could actually be used worldwide. No mean feat, but it did bring the experts together. Iris created another web application, which could be used to compile new instruments based on the D-score model. Child development experts were able to directly evaluate in that application how far the new instrument met certain statistical properties. Led by WHO, the Global Scales for Early Development (GSED) project took the

An accessible and affordable method

Tarun Dua, lead researcher on the GSED project at WHO: “When it comes to young children, we at WHO have always focused very much on their chances of survival. What we’ve seen in the last decade is that the focus has become broader: from surviving to thriving. In doing so, it’s extremely important to be able to properly measure child development and compare outcomes globally. We see the D-score as an accessible and affordable method that we believe can help us to monitor child development globally. However, more research is still needed in many different countries. Fortunately, there are currently large research teams at work in 12 countries. It is always difficult for people to change their way of thinking and working. That does make it a challenge. But if we all keep our eyes on what we are doing it for – i.e. better development opportunities for children around the world – we will get there. In doing so, it’s important to explain as clearly as possible what we want to do and what role the D-score can play in this. So, not in scientific terms, but in language that everyone understands. It should be clear to the general public and especially to policy officers and people working in public health. That way, we can make it perfectly clear that we all have the same goal.”

D-score as a basis to develop culturally sensitive measurement tools that are applicable worldwide. It just goes to show that using the GSED tools can help measure the D-score properly and, in turn, using the D-score can help compare the output of the GSED tools with that of other tools. That's the great thing about the D-score: it doesn't replace other instruments. The D-score works well with data retrieved by different validated instruments. It is a user-friendly method that can still produce clear, reliable scores for different data sets.

The expected impact of the D-score

While it is still too early to say whether the D-score will have a major global impact, it is interesting to note that WHO is currently working on a study to develop norms and standards for the D-score in 12 countries worldwide. The global impact of the D-score will certainly get a boost from that, as it did after WHO published standards on child growth in 2006. Once these standards became widely adopted, it then became possible to compare the growth of children worldwide. International organisations like WHO, Unicef, BMGF, and USAID have since used those standards to set informed priorities and can now estimate the impact of interventions. The D-score can perform a similar function for child development. Clear and comparable measurement helps raise awareness of backlogs and may lead a country or region to work more actively on that problem itself, or seek help to do so. Tracking the D-score over time provides insight into the effect of interventions, so that adjustments can be made where necessary. This is good news for the millions of children around the world who are unable to achieve their developmental potential because of the circumstances they are in.

Clear and comparable measurement helps raise awareness of backlogs and may lead a country or region to work more actively on that problem itself, or seek help to do so.

Visualising differences at local level

Laurens van Buren is a trained epidemiologist and has now been working as a public health researcher at the municipality of Utrecht for 16 years. He always found it difficult to build a picture of how young children in his community were developing because there was very little relevant data available. About 6 years ago, he wondered whether the Van Wiechen scheme, a standard part of youth healthcare assessments, might help here. Could that data also interest policymakers? No, he swiftly concluded. The data was too fragmented to be of any use.

During his research, he also came across TNO's D-score, which piqued his interest immediately. So, he decided to get in touch. This culminated in a close collaboration with the TNO team working on the D-score. Laurens was given the opportunity to make use of the D-score algorithm and run it on the Van Wiechen data in Utrecht. This results yielded some interesting insights. It became clear that children of parents with just primary or secondary education were 2.5 times more likely to be among the 10% slowest-developed children than children of parents with college or university education. No big surprise admittedly, but the figures show these children could use some extra support. The D-score results can also reveal differences at local area level. Now, that is information that could interest policymakers. "That kind of information enables us to take a local approach," Laurens points out. "I recently shared an overview of D-scores in various areas of Utrecht to our child development policy officer, and she found it extremely interesting."

The importance of a good team

Daring to deviate from the norm, launching projects to learn from, and – more than anything – persevering. That is what Stef van Buuren credits with helping to get the D-score off the ground step by step. The same goes for the scientific publications on the methodology, because while the initial articles were not picked up, there is now a great deal of interest in the literature. To put the theory into practice, the golden opportunity came in the form of the special software developed to make it easy to work with the D-score. Finally, it bears mentioning that besides Stef van Buuren and Iris Eekhout, there were many others at TNO who helped make the D-score a success over the years: Gert Jacobusse, Paul Verkerk, Elise Dusseldorp (currently Professor of Methods and Techniques at Leiden University), Hedwig Hofstetter, and Paula van Dommelen.

While we have done our best to paint the best possible picture of those who contributed what to the success of this case study, we might not have captured the full story. So, we close with some words of thanks. A big, big thank you to all those at TNO and elsewhere who did their bit and helped make this success story possible.

Publication details

Issue
TNO

Text
Menno de Boer

Editor
Wieke Ringeling

With thanks to:
Laurens van Buren - Public health researcher, Municipality of Utrecht
Stef van Buuren - Principal Scientist, TNO
Tarun Dua - Lead researcher on the GSED project, WHO
Iris Eekhout - Statistician, TNO
JGZ Kennemerland (photography)

TNO innovation for life

tno.nl